

(19) 世界知的所有権機関
国際事務局(43) 国際公開日
2005 年11 月17 日 (17.11.2005)

PCT

(10) 国際公開番号
WO 2005/109401 A1

- (51) 国際特許分類⁷: G10L 19/00
- (21) 国際出願番号: PCT/JP2005/008495
- (22) 国際出願日: 2005 年5 月10 日 (10.05.2005)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願2004-140117 2004 年5 月10 日 (10.05.2004) JP
- (71) 出願人 (米国を除く全ての指定国について): 日本電信電話株式会社 (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) [JP/JP]; 〒1008116 東京都千代田区大手町二丁目3 番1 号 Tokyo (JP).
- (72) 発明者; および
- (75) 発明者/出願人 (米国についてのみ): 大室 伸

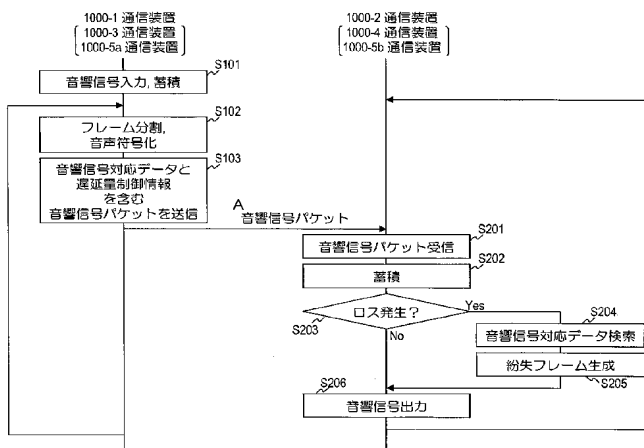
(OHMURO, Hitoshi) [JP/JP]; 〒1808585 東京都武蔵野市緑町三丁目9 番1 1 号 NTT 知的財産センタ内 Tokyo (JP). 森 岳至 (MORI, Takeshi) [JP/JP]; 〒1808585 東京都武蔵野市緑町三丁目9 番1 1 号 NTT 知的財産センタ内 Tokyo (JP). 日和▲崎▼ 祐介 (HIWASAKI, Yusuke) [JP/JP]; 〒1808585 東京都武蔵野市緑町三丁目9 番1 1 号 NTT 知的財産センタ内 Tokyo (JP). 片岡 章俊 (KATAOKA, Akitoshi) [JP/JP]; 〒1808585 東京都武蔵野市緑町三丁目9 番1 1 号 NTT 知的財産センタ内 Tokyo (JP).

- (74) 代理人: 中尾 直樹, 外 (NAKAO, Naoki et al.); 〒1600022 東京都新宿区新宿三丁目1 番2 2 号 新宿 NSOビル 4 階 Tokyo (JP).
- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM,

[続葉有]

(54) Title: ACOUSTIC SIGNAL PACKET COMMUNICATION METHOD, TRANSMISSION METHOD, RECEPTION METHOD, AND DEVICE AND PROGRAM THEREOF

(54) 発明の名称: 音響信号のパケット通信方法、送信方法、受信方法、これらの装置およびプログラム



1000-1... COMMUNICATION DEVICE
1000-3... COMMUNICATION DEVICE
1000-5a... COMMUNICATION DEVICE
1000-2... COMMUNICATION DEVICE
1000-4... COMMUNICATION DEVICE
1000-5b... COMMUNICATION DEVICE
S101... INPUT AND ACCUMULATE ACOUSTIC SIGNAL
S102... DIVIDE FRAME AND ENCODE AUDIO
S103... TRANSMIT ACOUSTIC SIGNAL PACKET CONTAINING ACOUSTIC SIGNAL-CORRESPONDING DATA AND DELAY AMOUNT CONTROL INFORMATION
A... ACOUSTIC SIGNAL PACKET
S201... RECEIVE ACOUSTIC SIGNAL PACKET
S202... ACCUMULATION
S203... LOS OCCURRED?
S204... SEARCH ACOUSTIC SIGNAL-CORRESPONDING DATA
S205... GENERATE LOST FRAME
S206... OUTPUT ACOUSTIC SIGNAL

(57) Abstract: When communicating an acoustic signal packet via an IP communication network, data corresponding to an acoustic signal (acoustic signal-corresponding data) is built in a packet different from the acoustic signal and transmitted. However, conventionally it is necessary to decide in advance in which packet the acoustic signal is built in and it has been impossible to perform dynamic modification. In this invention, a delay amount of the acoustic signal-corresponding data for the acoustic signal is built in the acoustic signal packet as delay amount control information. Moreover, the number of packets which have lost the communication network state in the burst loss detects fluctuation and acquires the number of packets to be accumulated at the reception side and the delay amount.

(57) 要約: 音響信号パケットを IP 通信網で通信する場合に、音響信号に対応するデータ (音響信号対応データ) を音響信号と異なるパケットに組み込んで送信している。しかし、従来はどのパケットに組み込むのかを事前に決めておく必要があり、動的に変更することもできなかった。本発明では、音響信号に対する音響信号対応データの遅延量を、遅延量制御情報として音響信号パケットに組み込む。また、通信網の状態をバーストロスで紛失したパケットの数がゆらぎを検出し、受信側で蓄積するパケットの数と遅延量とを求め



DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:

— 国際調査報告書

(84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE,

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

明 細 書

音響信号の packets 通信方法、送信方法、受信方法、これらの装置およびプログラム

技術分野

[0001] この発明は、デジタル化された音声・音楽などの音響信号をインターネットをはじめとする packets 通信網を介して送信する際に、特に packets 紛失対策をした通信方法、送信方法、受信方法、これらの装置及びプログラムに関する。

背景技術

[0002] 音声信号をボイスオーバ (Voice over) IP (インターネットプロトコル) 技術を利用して送信するサービスが普及しつつある。図1に示すように入力端子11よりの音声信号を音声信号送信部12で音声 packets に変換して IP 網をはじめとする packets 通信網13によって音声信号受信部14へ送信し、音声信号受信部14により音声信号を再生して出力端子15へ出力する。これをリアルタイム通信する場合、packets 通信網13の状態によっては通信網の途中において packets ロス (紛失) が生じ、それによって再生音声途切れるといった品質劣化が問題となっている。特に、インターネットなどのベストエフォートと呼ばれる通信サービスの場合には、packets ロスを許容しているため通信網の混雑時に特にこの問題が顕著である。

[0003] そこで、音声信号を packets 通信網で通信する場合には、packets ロスコンシールメントと呼ばれる手法を用いる。この手法は、packets が通信路の途中で消失した場合や通信路の遅延によって制限時間内に受信側に届かなかった場合に、消失または届かなかった packets (以下、「ロス packets 」又は「紛失 packets 」という。) に対応する区間の音声信号を受信側で推定して補償する方法が用いられる。図2は、図1における音声信号送信部12の一般的な構成例である。入力音声は入力バッファ21に蓄えられ、音声 packets 化部22で音声信号をフレームと呼ばれる一定の時間ごとに区切って音声 packets を生成し、packets 送出部23より packets 通信網に音声 packets を送出する。1フレームの時間長は一般には、10ミリ秒から20ミリ秒程度とすることが多い。

[0004] 図3は、図1における音声信号受信部14の一般的な構成例である。パケット通信網からパケット受信部31で受信した音声パケットは、ゆらぎ吸収バッファとも呼ばれる受信バッファ32に蓄えられる。正しくパケットが受信されたフレームについては、受信バッファから音声パケットが取り出され、音声パケット復号部33で音声信号に復号され、パケットロスしたフレームについては、紛失信号生成部34でパケットロスコンシールメント処理を行って音声信号が生成され、生成された音声信号が出力される。パケットロスコンシールメントの処理に、ピッチ周期(音声の基本周波数に相当する時間軸上での長さ)の情報を利用する場合には、出力音声信号を出力音声バッファ35に蓄え、ピッチ抽出部36でピッチ分析し、得られたピッチ周期の値を紛失信号生成部34に供給する。紛失信号生成部34で生成された信号は切替スイッチ37を通じて出力端子15へ出力され、パケットロスがない場合は音声パケット復号部33よりの復号音声信号が切替スイッチ37を通じて出力端子15へ出力される。なお、双方向で音声通信を行う通信端末は、各端末に送信部と受信部の両方を具備する。パケットロスコンシールメントの代表的な方法としては、非特許文献1に示す方法がよく知られている。非特許文献1に示す方法では、音声のピッチ周期をパケットロスコンシールメントに利用している。

[0005] 図4に、非特許文献1に示す方法でも用いられている、一般的なパケットロスコンシールメントの手法を示す。この図は、受信側の現在のフレームをフレーム n として、フレーム n に対応するパケットがロス(紛失)した場合の処理方法を示すものである。過去のフレーム(フレーム $n-1$ まで)の音声信号は、正しく復号されている、もしくはパケットロスがあった場合にもパケットロスコンシールメントの手法によって音声信号が生成されているものとする。フレーム n では音声パケットが受信できていないため、直前のフレーム $n-1$ の最後のサンプル点から、1ピッチ周期分の区間3Aの音声信号波形を切り出し、フレーム n の区間に切り出した1ピッチ波形を順にフレームの長さになるまで並べる(区間3B~3D)。

[0006] このように直前フレームの1ピッチ波形を並べる処理によってパケットロスとなったフレームの波形を生成すると、何も処理しないで0値でフレーム n の全サンプル点を埋めるのに比べて、自然な音質で再生することが可能である。

なお、このとき、1ピッチ波形を単純に並べた場合、接続点で波形が不連続となって、プツプツといった耳障りな音が生じることがある。その場合には、図5に示すような手法で接続点での不連続を防ぐ。図5では、説明をわかりやすくするために、フレーム n の区間を、切り出し波形ごとに段をずらして表示している。まず、フレーム $n-1$ の最後のサンプル点から、1ピッチ周期よりも少し長い、例えばピッチ長を L とすると4分の $5L$ ($5 \times L / 4$)の区間4Aの波形を切り出し、1ピッチ長ずつずらした4B, 4C, 4Dの位置に切り出した波形をそれぞれ並べる。このとき、切り出した波形は1ピッチ長よりも長いため、重なり区間4AB, 4BC, 4CDができる。これらの重なり部分は、例えば図6に示す三角窓関数をかけて重ね合わせるといった方法がとられる。図6は、横軸が時刻、縦軸が重みの値を示しており、 t_1 が重なり区間の始点、 t_2 が重なり区間の終点を示す。例えば図5における重なり区間4BCでは、重なり区間4BCの区間4B側の波形に重み関数 W_1 を乗じ、区間4C側の波形に重み関数 W_2 を乗じて加算することにより、区間BとCの両切り出し波形をなめらかにつなげることができる。このような重ね合わせの詳細についても、非特許文献1に記載されている。

[0007] パケットロスの発生する通信環境で非特許文献1に示す方法を利用した場合の音声品質は一般に良好と言われている。しかし、パケットロスが音声の子音と母音の境界付近で発生した場合に、耳障りな雑音が発生することがある(課題1)。また、連続した複数のフレームのパケットロスが発生(バースト的なロス(紛失)という)した場合、すなわち例えばフレーム長が20ミリ秒のときにパケットが2~3フレーム以上連続してロスした場合や、フレーム長が長い音声符号化方式で1フレームのパケットロスが発生した場合、すなわち例えばフレーム長が40ミリ秒や60ミリ秒の音声符号化方式のときにパケットロスが発生した場合に、耳障りなブザー音が発生することや、不自然な音声再生されるといった問題がある(課題2)。

[0008] 前記課題1の原因は、非特許文献1に示す方法がロスフレームの音声を生成するための手法として、直前のフレームの音声波形と基本的に特性が同じ波形を作成していることに起因する。即ち、子音と母音の境界付近の母音に近いフレームでフレームロスが発生すると、ロスしたフレームは母音区間であるにもかかわらず、子音と同じ特性の音声波形が生成される。また、母音から無音や子音へ変化する時刻でも同様

の雑音が発生する。

前記課題2は、パケットロスの発生した区間が子音と母音の境界付近でないときにも発生する。その原因は、パケットロスフレームの音声をさらに(自己回帰的に)利用して、連続した後続のフレームロス区間でも同じ特性の音声波形を生成するため、40ミリ秒～60ミリ秒以上の長い時間にわたって、同じ特性の音声波形が連続して再生されることに起因する。実際の音声はピッチ周期やパワが微小に変化しており、連続して同じ特性の音が再生されると、それは音声とは異なった音に聞こえる。

- [0009] 前記課題1、2を解決する手法として、非特許文献2の提案がある。この文献では、あらかじめk番目のフレームの補助情報をk+1番目のフレームに埋め込む。そして、k番目のフレームがパケットロスの発生によって届かなかった場合、k+1番目のフレームに埋め込まれた補助情報を、k番目のフレームのエラーを隠蔽するために用いている。

非特許文献1:ITU-T Recommendation G.711 Appendix I, “A high quality low-complexity algorithm for packet loss concealment with G.711”, pp.1-18, 1999.

非特許文献2:青木直史, “ピッチ波形複製法に基づくステガノグラフィを用いたVoIPにおけるパケット損失の一隠蔽法”, 電子情報通信学会誌B, Vol.J86-B, No.12, pp.2551-2560, 2003.

発明の開示

発明が解決しようとする課題

- [0010] パケットロスの発生する通信環境において、非特許文献1に示す方法のパケットロスコンシールメント手法を用いることは、再生音声の品質劣化を少なくする点で一定の効果がある。しかしながら、子音と母音の境界付近のフレームでパケットロスが発生した場合には、耳障りな雑音を十分に防ぐことができない、また連続した複数のフレームでパケットロスが発生した場合の再生品質が不自然である。

また、非特許文献2では、補助情報を追加することによるエラーの隠蔽方法については示されている。一方、補助情報を音響信号のフレームとどの程度ずらしたフレームと一緒に送信するのか(遅延量)については示されていない。実際の通信網では、通信状態は変化するため、パケットのゆらぎやバーストロスの発生頻度なども変化する

る。したがって、最適な遅延量も変化するはずである。しかし、従来技術には遅延量を動的に変化させる方法について示された例がない。また、対向する装置間で事前に遅延量を定めておく必要がある。

[0011] この発明はこれらの問題を解決し、より安定した品質で音響通信のできる方法、その装置、およびプログラムを提供することを目的とする。

課題を解決するための手段

[0012] この発明では、送信側で、音響信号と同一のパケットに、遅延量制御情報で指定する値だけフレーム番号が異なるフレームの音響信号に対応するデータ(以下、「音響信号対応データ」という。)と、遅延量制御情報とを組み込んで送信する。受信側では、受信できなかったパケットがある場合に、紛失した音響信号のフレーム(以下、「紛失フレーム」という。)と同じフレーム番号の音響信号対応データを、遅延量制御情報を用いて受信バッファ内のパケットから求める。そして、受信側では、求められた音響信号対応データを用いて紛失フレームの音響信号を生成する。

[0013] また、この発明では、遅延量を通信網の状態にあわせて変化させる。具体的には、受信側で、あらかじめ定めた期間ごとのパケットのゆらぎ時間の最大値、またはパケットロスが連続して発生するフレーム数(例えば、単独フレームでのロスは1、3フレーム連続してパケットロスが発生したときは3)の最大値(以下、「連続パケットロス数」という。)を検出し、受信バッファに蓄積されたパケットの数(以下、「蓄積パケット数」という。)と遅延量を、

ゆらぎ時間に相当するフレーム数(端数切り上げ) \leq 遅延量 \leq 蓄積パケット数

または

連続パケットロス数 \leq 遅延量 \leq 蓄積パケット数

を満たすように、制御する。

[0014] さらに、この発明では、送信側で、各フレーム内の音響信号に対応する少なくともピッチ周期を音響特徴量として求め、遅延量制御情報として指定されるフレーム数だけ、フレーム番号が異なる音響信号と音響特徴量を同一のパケットに組み込んで送信する。受信側で、出力しようとするフレームの音響信号に対応するパケットが紛失した場合は、紛失パケットの音響信号に対応する音響特徴量を、受信バッファ内のパケッ

トから求め、上記紛失パケットの音響信号のフレーム(以下紛失フレームという)と直近したフレームの音響信号から、上記求めた音響特徴量に含まれるピッチ周期に対応する長さの波形を切り出し、その切り出した波形を上記ピッチ周期で繰り返し並べ、この並べた信号を用いて上記紛失パケットの音響信号を生成する。

発明の効果

- [0015] この発明によれば、遅延量制御情報をパケットに組み込んで送信するとともに、各フレームの音響信号対応データをその音響信号とは別のパケットで送信している。つまり、遅延量を送信側が変更しても、受信側では、音響信号と音響信号対応データとの関係が把握できる。したがって、あるフレームの音響信号がパケットロスにより失われても、その音響信号の音響信号対応データはパケットロスすることなく受信されとともに、送信された遅延量制御情報から容易に求めることができる。

また、音響信号対応データとして、その音響信号のピッチ周期を用いる場合は、このピッチ周期と対応する長さで音響信号波形を切り出し、並べて、紛失音響信号を生成している。したがって、1フレーム単位のランダムなパケットロスの場合でも、連続して複数フレームのパケットロスが発生した場合でも、パケットロスが発生しない場合に近い音響品質で音響信号が再生されるため、安定した音響通信が実現される。

- [0016] このように、パケットロスの発生頻度の高いパケット通信網で、リアルタイムの安定した音響通信ができる。また、パケット通信網はコストを抑えるために、ある程度のパケットロスを許容するように設計するのが一般的であり、この発明の利用によって、回線自体のパケットロス率が低い高品質ネットワークが不要となるため、ネットワークコスト低減にも効果がある。

また、遅延量を通信網の状態にあわせて制御することにより、通信網の状態が相対的に良い場合には、受信バッファ内の蓄積パケット数を減らして、音声パケットを受信してから実際に音声信号が再生されるまでの遅延時間を短縮する。通信網の状態が相対的に悪い場合には、受信バッファ内の蓄積パケット数を増やして、(音声パケットを受信してから実際に音声信号が再生されるまでの遅延時間は増加するが)バーストパケットロスが発生しても確実に音響信号対応データを取得して、品質劣化の少ない音声信号を再生することができる。

図面の簡単な説明

- [0017] [図1]音声信号を音声パケットに変換し、パケット通信網によって通信する概念図。
- [図2]図1中の音声信号送信部12の一般的な機能構成例を示すブロック図。
- [図3]図1中の音声信号受信部14の一般的な機能構成例を示すブロック図。
- [図4]一般的なパケットロスコンシールメントの手法による紛失信号の生成を説明するための波形図。
- [図5]図4においてピッチ波形を繰り返して波形を作成する際に不連続音が生じないようにするための重ね合わせ手法(オーバーラップアッド処理)を説明するための波形図。
- [図6]オーバーラップアッド処理における三角窓関数の例を示す図。
- [図7]この発明による音響信号送信装置100の機能構成例を示す図。
- [図8]この発明による音響信号受信装置200の機能構成例を示す図。
- [図9]この発明による音響信号パケット送信方法の処理手順の例を示す図。
- [図10]この発明による音響信号パケット受信方法の処理手順の例を示す図。
- [図11]パケットの構成例を示す図。
- [図12]図8中の紛失信号生成部224の動作の一例を説明するための波形図。
- [図13]図8中の紛失信号生成部224の具体的機能構成例を示す図。
- [図14]紛失フレームの音響信号の特徴量の選択手順の例を示す図。
- [図15]連続してパケットロスが発生した場合の後ろ向き波形外挿処理を説明するための波形図。
- [図16]後ろ向き波形外挿のオーバーラップアッド処理を説明するための波形図。
- [図17]この発明による音声信号送信装置100-2の機能構成例を示す図。
- [図18]この発明による音声信号受信装置200-2の機能構成例を示す図。
- [図19]ゆらぎ時間と蓄積パケット数の対応の例を示す図。
- [図20]この発明による音声信号受信装置200-3の機能構成例を示す図。
- [図21]この発明による音声信号受信装置200-4の機能構成例を示す図。
- [図22]この発明による音声信号受信装置200-5の機能構成例を示す図。
- [図23]この発明による音声信号受信装置200-6の機能構成例を示す図。

[図24]この発明による音声信号受信装置200-7の機能構成例を示す図。

[図25]一方向に音声パケットを送る場合のシステム構成例を示す図。

[図26]音声パケットを一方向で送り、制御信号を逆方向に送る場合のシステム構成例を示す図。

[図27]送信装置300の機能構成例を示す図。

[図28]受信装置400の機能構成例を示す図。

[図29]双方向で音声パケットを送る場合のシステム構成例を示す図。

[図30]音声パケットの受信側から送信側への制御情報の送信を行わない場合の処理フロー例を示す図。

[図31]音声パケットの受信側から送信側への制御信号の送信を行う場合の処理フローの第1の例を示す図。

[図32]音声パケットの受信側から送信側への制御信号の送信を行う場合の処理フローの第2の例を示す図。

[図33]音声パケットの受信側から送信側への制御信号の送信を行う場合の処理フローの第3の例を示す図。

[図34]音声パケットの受信側から送信側への制御信号の送信を行う場合の処理フローの第4の例を示す図。

[図35]通信網の状態がパケット通信に与える影響は、双方向とも同じであると仮定した場合の処理フロー例を示す図。

発明を実施するための最良の形態

[0018] この発明は、コンピュータ本体とコンピュータプログラムとして実行することが可能であるし、デジタルシグナルプロセッサや専用LSIに実装して実現することも可能である。この発明は音声・音楽などの音響信号に適用できるが、以下では、図1に示したVoice over IPによる音声通信システムにこの発明を適用した場合について説明する。また、重複説明を避けるため、同じ機能を備える構成部や、同じ処理をする過程には同じ番号を付している。

以下では、実施例1として、音響信号対応データとしてピッチ周期を用いた場合の音声信号送信装置と音声信号受信装置の機能構成や処理フローを説明する。実施

例2では、遅延量と蓄積パケット数を通信網の状態に連動させる方法について説明する。なお、実施例2で用いる音響信号対応データは、ピッチ周期に限る必要はなく、音響信号そのものでもよく、また他の音響信号特徴量でもよい。

[実施例1]

音声信号送信装置100(図1中の送信部12と対応)の機能構成例を図7に、音声信号受信装置200(図1中の受信部14と対応)の機能構成例を図8にそれぞれ示し、音声信号送信装置100の処理手順の例を図9に、音声信号受信装置200の処理手順を図10にそれぞれ示す。

送信側

送信装置100において、入力音声信号は入力バッファ111に蓄えられ、音声信号をフレームと呼ばれる一定の時間ごとに区切って、つまりフレーム分割して(ステップS1)音声波形符号化部112に送る。1フレームの時間長は一般には、10ミリ秒から20ミリ秒程度とすることが多い。

[0019] 音声波形符号化部112では、例えば、ITU-Tの標準符号化方式であるG. 711(μ -law PCMまたはA-law PCM方式)を適用して入力音声信号を音声符号に変換する(ステップS2)。なお、G. 711以外の標準方式や、非標準の音声符号化方式など、符号化方式は任意の方式を用いてよい。音声符号はパケット構成部113に送られる。

音声特徴量計算部114では、入力バッファ111に蓄えられた音声信号を用いて、当該フレームにおけるその音声信号の音声特徴量を計算する(ステップS3)。音声特徴量とは、ピッチ周期(音声の基本周波数に相当)、パワなどを指し、これらの特徴量のいずれかのみを利用することもあるし、全部を利用することもある。ピッチ周期はピッチ部114aで例えば、音声信号波形または音声信号波形にスペクトル包絡の逆特性を持つフィルタをかけた信号の自己相関係数を計算することによって得られる(ステップS3a)。パワは、パワ部114bでフレーム内の音声信号の全サンプルの二乗和により計算する(ステップS3b)。この二乗和をフレーム長のサンプル数(音声信号のサンプリング周波数が8kHzでフレーム長が20ミリ秒の場合は160)で除して平方根をとった値(1サンプルあたりのパワの平方根)をパワを表す特徴量として利用して

もよい。

- [0020] 音声特徴量符号化部115では、前記音声特徴量を定められたビット数(またはバイト数)で表すことができるように量子化した後、符号に変換する(ステップS4)。例えば、ピッチ周期の場合、音声のサンプリング周波数が8kHzサンプリングで、ピッチ周期の探索範囲が40サンプル～120サンプルの場合には、符号として0～80を割り当てることができる、7ビットで符号化することができる。探索範囲が20サンプル～160サンプルの場合は符号として0～140を割り当てるので8ビットである。パワの場合は、前記1サンプルあたりのパワの平方根にG. 711方式を適用すれば8ビットで符号化することができる。(パワに負の値はないので、実際には正負の符号ビットを除いた7ビットでよい。)

符号化された音声特徴量は、シフトバッファ116に送られる。シフトバッファ116では、あらかじめ指定された数のフレームにわたって音声特徴量の符号を保持する。後述する遅延量制御情報が端子117から入力されると、遅延量制御情報で指定されたフレーム数だけ前、つまり過去のフレームの音声信号の音声特徴量の符号(「補助情報」ともいう。)をパケット構成部113に送る。例えば、現在のフレームをn、遅延量制御情報が3とすると、フレームn-3で生成された音声特徴量符号がシフトバッファ116で3フレーム遅延されてパケット構成部113へ送られる(ステップS5)。

- [0021] 端子118から後述するバッファ残量が入力され、バッファ残量符号化部119で符号化される(ステップS6)。そのバッファ残量符号もパケット構成部113に送られる。パケット構成部113では、前記音声信号波形を符号化した符号と、音声特徴量の符号と、遅延量制御情報と、バッファ残量符号を用いてパケットを構成する(ステップS7)。なお、遅延量制御情報とバッファ残量符号はパケットに組み入れない場合もある。この点については後述する。

パケット送出部121は、パケット構成部113で作成されたパケットの情報を受け取り、音声パケットとしてパケット通信網に送出する(ステップS8)。

- [0022] パケット構成部113での、パケットの構成例を図11に示す。一般にパケットはヘッダ領域41とデータ領域42からなり、ヘッダ領域41の中でも必須領域41aとオプション領域41bがある。パケットの宛先やフレーム番号などのタイムスタンプはヘッダ領域4

1に格納される。音声信号波形を符号化した符号データはデータ領域42に格納される。音声特徴量の符号は、波形データに比べてサイズ(バイト数)が少ないため、ヘッダ領域41のオプション領域HOP41bに格納する方法と、データ領域42の先頭D1または末尾DEに格納する方法のいずれでもよい。オプション領域HOPに格納する方法が、この発明を適用しない方法との互換性を維持しやすいという長所がある。また、パケットのヘッダ領域41は4バイトの倍数とするのが一般的である。音声特徴量として、ピッチ周期とパワをそれぞれ1バイトずつで量子化して送る場合(7ビットの場合は1バイト=8ビットに切り上げる)、ピッチ周期とパワ情報の合計2バイトに、遅延量制御情報とバッファ残量符号を各1バイトで表現して合計4バイトとなる。これらを、例えばオプション領域HOP41bの4つの領域OP1, OP2, OP3及びOP4に各1バイトずつ格納する。遅延量制御情報を補助情報(音声特徴量符号)とともに送るのは、補助情報がどのフレームに対応する補助情報であるかを、受信側で知るためであり、フレーム n を基準とした相対的なタイムスタンプとみることができる。また、バッファ残量符号を補助情報とともに送るのは、相手側が自分宛にパケットを送るときに、補助情報をフレーム n と相対的に何フレームずれて送ってもらうのが適当であるかを伝えるためである。

- [0023] 補助情報がフレーム n と相対的に何フレームずれたものであるかをフレーム毎にダイナミックに変動させない場合は、遅延量制御情報やバッファ残量符号を補助情報とともに送ることは不要であり、図7中のバッファ残量符号化部119、図8中のバッファ残量復号部216、受信バッファ残量判定部218、図9中のステップS6は省略される。あらかじめ決められた相対的なフレームのずれを送信側と受信側で事前に(例えば呼の接続時のネゴシエーションで)知っていればよいからである。その場合には、遅延量制御情報とバッファ残量符号を各1バイトで送る必要がなく、ピッチ周期とパワ情報の合計2バイトだけで済む。ヘッダ領域が4バイト単位であるので、補助情報を2フレーム分、例えば、フレーム $n-3$ とフレーム $n-6$ の各音声信号に対する各補助情報をフレーム n のパケットにつけて送ることができる。したがって、音声信号符号とその補助情報のいずれもがパケットロスとなる確率が低くなる。またこの遅延量制御情報を一定にする場合は、音声信号符号と、その補助情報が共にパケットロスにならないよ

うに、つまりバースト的ロスに耐え、かつリアルタイム性が失われない程度に、60ミリ秒、80ミリ秒あるいは100ミリ秒程度と対応するフレーム数とする。

受信側

図8の音声信号受信装置200において、パケット受信部211は、パケット通信網から音声パケットを受信し(ステップS21)、受信バッファ212に蓄積する(ステップS22)。受信バッファ212はゆらぎ吸収バッファとも呼ばれる。前述したように、音声パケットには、音声信号波形を符号化した符号と、ピッチ、パワ等の符号の補助情報、遅延量制御情報、バッファ残量符号が含まれている。補助情報がフレームnと相対的に何フレームずれたものであるかをフレーム毎にダイナミックに変動させない場合は、遅延量制御情報とバッファ残量符号は不要である。ただし、以下は遅延量制御情報とバッファ残量符号がともにパケットに組み込まれていることを前提にして説明する。

[0024] 受信した音声パケットに含まれる音声信号波形を符号化した符号は、音声パケット復号部213に送られ、音声信号波形に復号される(ステップS24)。パケットロスが発生していないフレームでは、音声パケット復号部213の出力信号が切替スイッチ214を通じて出力端子215に再生音声として出力される(ステップS36)。

バッファ残量復号部216では、受信した音声パケットに含まれるバッファ残量符号から、補助情報を何フレームまでずらしてパケットにつけるかを指定する遅延量制御情報を求める。求められた遅延量制御情報は、図7中の端子117、つまりシフトバッファ116とパケット構成部113に送られる。相手側バッファ残量と遅延量制御情報の関係については後述する。

[0025] 受信した音声パケットに含まれる遅延量制御情報は、紛失処理制御部217で利用される。紛失処理制御部217における処理は後で詳細に述べる。

受信バッファ残量判定部218は、受信バッファ212に蓄積されているパケットのフレーム数を検出する。例えば、フレームnが音声パケット復号部213で復号されているときに、受信バッファ212にフレームn+1, n+2, n+3の音声信号符号のパケットが蓄積されている状態であれば、バッファ残量、つまり蓄積パケットのフレーム数は3である。このバッファ残量は、図7中の端子118、つまりバッファ残量符号化部119に送られる。なお、図7に示す送信装置100と、図8に示す受信装置200が相互に連携す

るこの通信方法は、双方向音声通信を行い、各端末が送信装置100及び受信装置200の両方を具備していることが前提である。

- [0026] なお、音声信号送信装置100の端子117は、送信部117自体を制御するための信号の入力端子である。つまり、入力端子117に入力された遅延量制御情報にしたがって、音声特徴量が遅れて送信されるパケットに組み込まれる。一方、端子118は、通信先の音声信号受信装置200に送信するための情報の入力端子であり、この情報によって、音声信号送信装置100は制御されない。つまり、端子118に入力された情報は、通信先に送信されるだけである。

紛失検出部219はパケットロス(紛失)を検出する(ステップS23)。パケット受信部211で受信されたパケットはそのパケット番号にしたがって、つまりフレーム番号の順に受信バッファ212に格納される。例えば、受信したパケットのフレーム番号がその直前に受信したパケットのフレーム番号より3つ多ければ、直前に受信したパケットの格納位置から2つのパケットを格納する位置をあけてその受信パケットを格納する。受信バッファ212からはその格納位置の順にパケットが読み出されるが、読み出される際に読み出す格納位置にパケットがない場合には、その読み出し動作の直前に、パケットロス(紛失)が発生したと紛失検出部219が判定し、切替スイッチ214を紛失処理制御部217の出力側に切替える。紛失処理制御部217におけるパケットロスコンシールメント制御処理の詳細を説明する。

- [0027] フレーム n において、パケットロスが発生したとする。パケットロスが発生した場合は、受信バッファ探索部221が受信バッファ212に蓄積された受信パケットを探索し、フレーム $n+1$ 以降で受信されているパケットの中で、紛失フレーム n に時間の近い(タイムスタンプが直近の)パケットを探索する(ステップS25)。この探索の結果フレーム $n+i$ のパケットが得られたとする。例えば、フレーム $n+1$ がロス(紛失)していないときは $i=1$ 、フレーム $n+1$ 、フレーム $n+2$ とも連続してロスしてフレーム $n+3$ がロスでない場合は $i=3$ といった具合である。フレーム $n+i$ のパケットに含まれる音声信号波形を符号化した符号は、先読み音声波形復号部222で復号され(ステップS26)、フレーム $n+i$ の音声信号波形が得られる。さらに受信バッファ探索部221において受信バッファ212に蓄積されたパケットの中から、紛失フレーム n の音声信号に対応する

補助情報が付加されたパケットを探索する(ステップS27)。

- [0028] 前述したように、遅延量制御情報(相対的なタイムスタンプ)がパケットに付加されている場合には、その遅延量制御情報をもとにフレーム n の音声信号と対応する補助情報が付加されたパケットを探索することができる。例えば、フレーム $n+3$ のパケットに遅延量制御情報として3が付加されていれば、それはフレーム n の音声信号と対応する補助情報である。補助情報が付加されるパケットはフレーム n と相対的に何フレームずれたものであるかをフレーム毎にダイナミックに変動させない場合は、遅延量制御情報を送信側でパケットに付加する必要がなく、その場合は、あらかじめ決められた遅延量を前提として紛失フレーム n の音声信号と対応する補助符号が付加されたパケットを探索する。例えば、あらかじめ決められた遅延量が5のとき、紛失フレーム n の音声信号対応補助情報は、フレーム $n+5$ のパケットに付加されている。前述したように1フレーム分のパケットに2フレーム分の補助情報を付加している場合は、探索の結果、紛失フレーム n の音声信号対応補助情報が付加されたパケットは受信バッファ212内に2つ見つかることがあるので、その場合はいずれかを利用すればよい。

- [0029] このパケット探索で見つかれば(ステップS28)、探索された紛失フレーム n の音声信号に対応する補助情報は、音声特徴量復号部223において紛失フレーム n の音声信号のピッチ情報とパワ情報に復号され(ステップS29)、紛失信号生成部224に送られる。

一方、非特許文献1に示す方法と同様に、出力端子215に出力される出力音声信号は出力音声バッファ225に蓄えられ(ステップS36)、ステップS28でパケット検索により見つからなければ出力音声バッファ225の出力音声信号はピッチ抽出部226においてピッチ周期の分析が行われる(ステップS30)。ピッチ抽出部226において抽出されるピッチは、紛失フレームの直前フレーム $n-1$ の音声信号に対応するピッチである。直前フレーム $n-1$ の音声信号に対応するピッチは、紛失信号生成部224に送られる。

- [0030] 紛失信号生成部224の動作イメージを図12に、機能構成例を図13に示す。フレーム n を現在のフレームとし、フレーム n に対応するパケットがロス(紛失)したものとする

。非特許文献1に示す方法は、パケットロスの発生した直前のフレーム $n-1$ の波形をピッチ単位で繰り返すもので、ピッチ周期、パワを含む信号の特性は直前のフレームと同一であった。この実施形態では図13に示すように、直前のフレーム $n-1$ の信号からピッチ周期と対応する信号波形を切り出し、この信号波形を、非特許文献1に示す方法と同様に、前向き波形外挿部51においてピッチ周期単位で波形を繰り返して1フレーム長の波形10-1を作成する(ステップS31)。このとき、非特許文献1に示す方法では、ピッチ周期として直前のフレーム $n-1$ における音声信号のピッチ長を用いているが、この発明では、補助情報を復号して得られた紛失フレーム n の音声信号のピッチ長を用いる。補助情報を組み込んだパケットも紛失した場合など、ステップS28でパケットを見つけることができず補助情報が得られていない場合は、ステップS30で得られた直前のフレーム $n-1$ の音声信号のピッチ長で代用する。

[0031] また、この実施例ではフレーム $n+1$ 以降で、受信バッファ212内に正常に受信できたフレーム $n+i$ の音声信号(つまり、フレーム $n+1$ のパケットが受信できていればフレーム $n+1$ の信号、フレーム $n+1$ のパケットも連続して紛失して、フレーム $n+2$ のパケットが受信できた場合にはフレーム $n+2$ の信号)からピッチ周期と対応する信号波形を切り出し、後ろ向き波形外挿部52でピッチ周期単位の波形を繰り返して1フレーム長の波形10-2を作成する(ステップS32)。このとき、時間軸上で未来の波形を使って外挿するため、波形の繰り返し方向が時間軸の負の向きであることに注意する。また後ろ向き波形外挿部52で使用するピッチ周期は、前向き波形外挿部51と同様に、補助情報を復号して得られた紛失フレーム n の音声信号のピッチ長とする。しかし、補助情報を組み込んだパケットもロスした場合など、補助情報が得られていない場合は、フレーム $n-1$ の音声信号のピッチ長で代用するか、または、フレーム $n+1$ の音声信号対応の補助情報が得られていれば、フレーム $n+1$ の音声信号のピッチ長で代用してもよい。

[0032] つまり紛失フレーム n の音声信号を生成しようとする場合に、波形外挿の際の切り出しピッチ長は例えば図14に示すように、紛失フレーム n の音声信号に対する補助情報が格納されたパケットがロスでなければ(ステップS41)、その補助情報を復号して用いる(ステップS42)。その補助情報が付加されたパケットがロスであれば、外挿が

前向きかを調べる(ステップS43)。前向きであれば、直前のフレーム $n-1$ の出力音声信号からピッチ長を分析して、そのピッチ長を用いる(ステップS44)。外挿が前向きでなければ、直後のフレーム $n+1$ の音声信号に対する補助情報が付加されたパケットを探索する。これがロスでなければ、ステップS42に移ってその補助情報を復号してピッチ長を用いる。なお、このフレーム $n+1$ の音声信号の補助情報も得られなければ、ステップS44に移る。前記前向き波形外挿部51と、後ろ向き波形外挿部52のそれぞれ処理イメージを図12に示す。9-1は過去の波形、9-2は先読みの波形を示している。前向き波形外挿では、フレーム $n-1$ の音声信号から切り出した1ピッチ長の波形10A1を区間10B1, 10C1, 10D1のように繰り返して波形10-1を作成する。後ろ向き波形外挿では、フレーム $n+1$ の音声信号から切り出した1ピッチ長の波形10A2を、区間10B2, 10C2, 10D2のように繰り返して波形10-2を作成する。図15の波形9-2に、フレーム n 、フレーム $n+1$ が連続してパケットロスになり、フレーム $n+2$ が正常に受信できた場合の後ろ向き波形外挿のイメージを示す。この場合は、フレーム $n+2$ の音声信号から1ピッチ波形13Aを切り出し、これをフレーム $n+1$ のフレーム $n+2$ 側からフレーム n のフレーム $n-1$ 側に向って区間13B, 13C, ..., 13Fと繰り返すことによってフレーム n の外挿波形10-2を得る。

[0033] なお、非特許文献1に示す方法の説明において図5で述べたように、単純に1ピッチ長に切り出した波形を並べると、接続点で不連続音が出る。そこで、前向き波形外挿では図5と同様に重なり区間を設けて窓掛け加算(Overlap add, OLA)を行うのがよい。同様に後ろ向き波形外挿においても図16に示すように、窓掛け加算(OLA)を、図5の時間軸を逆転した方法で実現できる。

波形10-1と波形10-2には、図13に示すように乗算部53と乗算部54においてそれぞれ重み $W1$ と $W2$ が乗算され、加算部55において加算され波形 y_n となる(ステップS34)。重み $W1$ と $W2$ はそれぞれ $1/2$ にする簡略化された方法と、波形10-1と波形10-2の「信頼度」に基づいて重み決定部56でそれぞれ重み $W1$ と $W2$ を決定する(ステップS34)方法がある。「信頼度」に基づくとは、過去最後に正常受信できたフレームと現在のフレーム n とのフレーム番号差と、受信バッファ212内から探索して得た現在よりも後の時刻(タイムスタンプ)の受信パケットのフレームと現在のフレ

ーム n とのフレーム番号差とのうち、番号差の小さいほうを信頼度が高いとみなすことである。このような「信頼度」を利用することは、バースト的なパケットロスの場合に有用である。具体例を以下にいくつか挙げる。

[0034] 例1:直前のフレーム $n-1$ ではパケットロスが発生しておらず、フレーム n でパケットロスが発生、直後のフレーム $n+1$ ではパケットロスが発生しなかった場合は、前向き外挿波形10-1と後ろ向き外挿波形10-2の「信頼度」は同じとみなされるので、重みはそれぞれ $1/2$ とする。

例2:直前フレーム $n-1$ ではパケットロスが発生しておらず、フレーム n でパケットロスが発生、直後フレーム $n+1$ もパケットロス、更に次のフレーム $n+2$ ではパケットロスが発生しなかった場合は、前向き外挿波形10-1と後ろ向き外挿波形10-2の「信頼度」は直前のフレーム $n-1$ から前向き波形外挿によって得られた波形10-1のほうが高いと推測される。したがって、前向き外挿波形10-1に乗ずる重みは $2/3$ 、後ろ向き外挿波形10-2に乗ずる重みは $1/3$ とする。これは正常にパケットが受信された前、後の直近フレーム $n-1$ と $n+2$ から音声信号を生成しようとするフレーム n までの「距離」が $1:2$ であるため、信頼度を $2:1$ とみなす。

[0035] 例3:フレーム $n-3$ が正常受信、フレーム $n-2$ 、フレーム $n-1$ が連続パケットロスの後、フレーム n でパケットロス、フレーム $n+1$ ではパケットロスが発生しなかった場合は、前向き外挿波形10-1より後ろ向き外挿波形10-2のほう信頼度が高いと推測される。前向き外挿波形10-1に乗ずる重みは $1/4$ 、後ろ向き外挿波形10-2に乗ずる重みは $3/4$ とする。これは正常にパケットが受信された前、後直近のフレームから音声信号を生成しようとするフレーム n までの「距離」が $3:1$ であるため、信頼度を $1:3$ とみなす。従って図13中の重み決定部56では、例えば正常にパケットが受信された前及び後のフレーム n_B 及び n_L と音声信号を生成しようとするフレーム n_N とのそれぞれの差 $\Delta n1$ 及び $\Delta n2$ が差回路56a及び56bで計算される。これらの差 $\Delta n1$ と $\Delta n2$ が和回路56cで加算される。この加算値により差 $\Delta n2$ を除算回路56dで割算して前向き外挿波形10-1に対する重み $W1 = \Delta n2 / (\Delta n1 + \Delta n2)$ を求める。また、前記加算値により差 $\Delta n1$ を除算回路56eで割算して後ろ向き外挿波形10-2に対する重み $W2 = \Delta n1 / (\Delta n1 + \Delta n2)$ を求める。

- [0036] 前記の例では、前向き外挿波形10-1の全サンプル点、後ろ向き外挿波形10-2の全サンプル点に対して、フレーム内で一律の重みを乗じた。しかし、正常にパケットが受信されたフレームからの「距離」をサンプル点ごとに決定して、サンプル点毎に信頼度に応じた重みを設定してもよい。

図13中のパワ補正部57は、加算波形 y_n のパワを、フレーム n の音声信号の補助情報を復号して得られたパワ情報によって補正し(ステップS35)、フレーム n の音声信号 x_n として出力端子215に出力する(ステップS36)。パワを補正するとは、加算波形 y_n のパワが、補助情報が示すパワと同等になるように波形 y_n の振幅を増減することである。ここで、同等とは、完全にパワが一致するか、または聴感上の影響を考慮して出力する音声信号 x_n のパワのほうが多少低くなる程度にすることをいう。

バッファ残量と遅延量制御情報

前述の相手側バッファ残量と遅延量制御情報の関係について説明する。

- [0037] 遅延量(=補助情報を音声信号のいくつずらしたフレームと一緒に送信するかのずれ量)は、大きければ大きいほど、バースト的なロスに強くできる。しかし、受信側で補助情報を利用してパケットロスコンシールメント(紛失パケットの音声信号の生成)の処理をするには、補助情報が到着するまで音声の再生ができない。したがって、音声パケットを受信してから実際に音声信号が再生されるまでの遅延(再生遅延)が増大する。また、双方向に通信する場合には、相手が発声してから自分側で音声再生されるまでの遅延(通話遅延)が増大する。遅延量が小さい場合には、再生遅延も小さくでき、単独の(ランダムな)パケットロスには耐えられる。しかし、バースト的なロスが発生すると、補助情報も失われる可能性が高い。つまり、遅延量と効果の間にはトレードオフが生じる。そのような場合に最適なのは、通信相手側の受信バッファに、その時点で何フレーム分のパケットが蓄積されているかという情報を相手側から受け取り、相手側受信バッファにおけるパケット蓄積フレーム数に相当する遅延量を設定することである。
- [0038] 前述したように、パケットの到着時間のゆらぎを吸収するために受信バッファには一定量のパケットが蓄積されるようにするのが一般的である。特に、パケット通信網の通信状態があまりよくない場合には、相手側は受信バッファに蓄積するパケットの数を

大きくして、パケットの到着ゆらぎによって音切れが生じないようにするのが一般的である。つまり、もともと受信バッファに蓄積されたパケットの数分の再生遅延が生じている。従って、相手側受信バッファに蓄積されたパケットのフレーム数(バッファ残量)に相当する値を遅延量として設定して補助情報を送信すれば、新たな再生遅延を増加させることなく、補助情報送信の効果を最大にすることができる。なお、相手側受信バッファに蓄積されたパケットのフレーム数以下の遅延量であれば、新たな再生遅延を増加させることがない。そこで、想定されるバースト的なパケットロスの特性に応じて、相手側受信バッファのパケット蓄積フレーム数以下の遅延量に設定してもよい。また、相手側受信バッファのバッファ残量が時々刻々変動することを考慮すれば、相手側バッファ残量よりも1〜2フレーム程度少ない遅延量に設定する方法も考えられる。

[0039] 音響特徴量としてはピッチ周期のみでもよい。この場合は図7中のパワ部114b、図9中のステップS3b、図10中のステップS35、図13中のパワ補正部57は省略される。また送信側で音響信号波形を符号化することなく、例えば入力されたPCM音響信号をそのままパケットに組み込んでもよい。この場合は、図7中の音声波形符号化部112、図8中の音声パケット復号部213、先読み音声波形復号部222、図9中のステップS2、図10中のステップS24、ステップS26は省略される。更に音響特徴量についても、これらは通常デジタル演算により求められるから、その計算結果が例えば8ビットで得られるようにすれば、特に音響特徴量も符号化する必要がなく、図7中の音声特徴量符号化部115、図8中の音声特徴量復号部223、図9中のステップS4、図10中のステップS29は省略される。

[0040] 紛失信号の生成において、前後のパケットが正常に受信され、フレームnの1パケットのみが紛失(ロス)した場合、前向き外挿波形10-1のみあるいは後ろ向き外挿波形のみを用いてもよい。またバースト的なパケットロスの場合は、正常に受信された直後のフレームの音響信号としては前向き外挿波形のみを、正常に受信された直前のフレームの音響信号に対しては、後ろ向き外挿波形のみを用いることとしてもよい。

図7に示した音響信号送信装置100、図8に示した音響信号受信装置200はそれぞれコンピュータにより機能させてもよい。その場合は、図9に示したパケット送信方法の処理手順又は図10に示したパケット受信方法の処理手順の各過程をコンピュ

ータに実行させるための音響信号パケット送信プログラム又は音響信号パケット受信プログラムをコンピュータにCD-ROM、磁気ディスク装置、半導体記憶装置などの記録媒体からインストールし、あるいは通信回線を介してダウンロードしてそのプログラムをコンピュータに実行させればよい。

[実施例2]

実施例1では、音声信号送信装置100と音声信号受信装置200について、音声特徴量を音声信号とは異なるパケットに組み込んで送信すること、音声信号に対する音声特徴量の遅延量を受信側のバッファ残量を用いて制御すること、パケットロスが生じた場合の紛失フレームの生成方法などについて説明した。本実施例では、バッファ残量以外の情報を用いて遅延量を制御する方法を説明する。また、音声信号と異なるパケットに組み込んで送る情報も、音声特徴量に限定せず、音響信号に対応するデータ(以下、「音響信号対応データ」という。)として説明する。パケットが紛失した場合の紛失フレームの生成方法としては、音響特徴量から生成する方法以外にも考えられるからである。例えば、最も簡単な方法としては、音響信号対応データとして、音響信号そのものとする方法がある(つまり、音響信号を2回送る。)。この場合、紛失フレームの生成とは、別のパケットで送られた紛失フレームと同じフレームの音響信号をそのまま使えばよい。

音声信号送信装置と音声信号受信装置

図17は、音声信号送信装置100-2の機能構成例を示している。図7の音声信号送信装置100との違いは、バッファ残量符号化部119の代わりに情報符号化部131を、音声特徴量計算部114の代わりに音響信号対応データ生成部132を、音声特徴量符号化部115の代わりに音響信号対応データ符号化部133を備えていることである。動作の手順は、図9に示した手順と同じであり、単純にバッファ残量を端子118で受け取る情報に、補助情報を音響信号対応データに変更するだけである。端子118は、対向する通信装置を制御するための情報を受け取る端子であり、当該音声信号送信装置100-2とセットで備えられる音声信号受信装置からの情報、または記録装置にあらかじめ定めた情報を受け取る。この情報としては、音声信号受信装置の受信バッファに蓄積されたパケット数(バッファ残量)、音声信号受信装置が受信した

パケットのゆらぎ情報、音声信号受信装置が検出したバーストロスで紛失したパケットの連続フレーム数(連続パケットロス数)などの情報、前記ゆらぎ情報や連続パケットロス数から求められる受信バッファに蓄積すべきパケット数(蓄積パケット数の目標値)、もしくは記録装置にあらかじめ定めておいた定数などが考えられるが、これらに限定されるものではない。これらの情報を総称して、遅延量制御情報の設定依頼情報と呼ぶことにする。

[0041] 図18は、音声信号受信装置200-2の機能構成例を示している。図8の音声信号受信装置200との違いは、バッファ残量復号部216の代わりに遅延量制御部341を備えていることと、受信バッファ残量判定部218で受信バッファ212のバッファ残量を判定する代わりに、パケット受信部211からゆらぎを検出し(ゆらぎ検出部232)、受信バッファ212に蓄積すべきパケット数(蓄積パケット数の目標値)を定めて、受信バッファ212に実際に蓄積されたパケット数(バッファ残量)を前記蓄積パケット数の目標値に近づけるように受信バッファ212を制御する(受信バッファ制御部231)ことである。遅延量制御部341は、通信先の音声信号送信装置から送られた通信先の遅延量制御情報の設定依頼情報をパケットから取り出す。さらに、遅延量制御部341は、当該音声信号受信装置200-2とセットで備えられる音声信号送信装置の遅延量を決め、端子117へ伝える。蓄積パケット数の目標値から遅延量を決める方法は、前記のバッファ残量から遅延量を決める方法のバッファ残量を蓄積パケット数の目標値に置き換えればよい。これは、蓄積パケット数の目標値を設定すると、蓄積パケット数の変更制御開始から完了までに一定の遷移時間を必要とするものの、一定時間経過後に、実際のバッファ残量が蓄積パケット数の目標値に一致することを前提としているためである。また、音声信号受信装置200-2とセットで備えられる音声信号送信装置へ、受信バッファ制御部231から送る情報は、バッファ残量の代わりに、蓄積パケット数の目標値である。実際のバッファ残量は蓄積パケット数の目標値に近づくように制御されるが、ゆらぎなどのネットワーク状態が変動しているときには、目標値の変更から実際のバッファ残量が目標値になるまでにある程度の時間を要する。そこで、ある時点ではバッファ残量は蓄積パケット数の目標値に等しく、ある時点では蓄積パケット数の目標値に対して実際のバッファ残量が過不足している状態が生じる。

なお、紛失状態判定部233を設け、パケットロスも考慮して蓄積パケット数を決定する方法もある。具体的には、紛失状態判定部233で、紛失検出部219で検出するパケットロスの情報も受け取り、バースト的なパケットロスで何パケットが連続して紛失したか(以下、「連続パケットロス数」という。)の最大値を求める。受信バッファ制御部231は、ゆらぎと連続パケットロス数の両方の情報から蓄積パケット数の目標値を決める。

[0042] 次に、ゆらぎから蓄積パケット数の目標値を決める方法を説明する。例えば、フレーム長が20ミリ秒の場合には、ゆらぎがなければパケットは20ミリ秒ごとに受信されるはずである。しかし、ネットワークの状態によって、あるパケットとあるパケットの受信間隔が40ミリ秒であったり、5ミリ秒であったり、時に100ミリ秒であったりする。そこで、ゆらぎ検出部232は、例えばパケットの受信間隔が毎フレーム20ミリ秒であれば、ゆらぎ時間は0ミリ秒と判断する。パケットの受信間隔が100ミリ秒のときは、ゆらぎ時間は80ミリ秒と判断する。パケットの受信間隔が5ミリ秒の場合は、おそらく直前のパケットの到着がゆらぎによって遅れたためと考えられるが、ゆらぎ時間がマイナス15ミリ秒、あるいは絶対値をとって15ミリ秒と判断する。このように時々刻々観測されたゆらぎ時間の時系列を用いて、フレームごと(例えば20ミリ秒ごと)に過去一定時間(例えば、2秒間)におけるゆらぎの最大値を求める。ゆらぎ時間の時系列から最大値を求める方法は、ゆらぎ時間を一定時間蓄積して、その中での最大値を検索するとよい。ゆらぎ時間の最大値は、受信バッファ制御部231に送られる。受信バッファ制御部231は、得られたゆらぎ時間の最大値に対応するフレーム数、すなわち、ゆらぎ時間の最大値をフレーム長で除した数(整数値に端数切り上げ)、あるいはそれに1を加えたフレーム数分を蓄積パケット数の目標値として、受信バッファ残量が目標値に近づくように制御する。また、図19に示すようなゆらぎ時間の最大値と蓄積パケット数の目標値の対応表によって、蓄積パケット数の目標値を決めてもよい。蓄積パケット数の目標値はフレームごとに決定されるが、ゆらぎ時間の最大値に変化がないときには、蓄積パケット数の目標値も同じ値が維持される。なお、前記過去一定時間におけるゆらぎ時間の最大値を求め、ゆらぎ時間の最大値に対応する蓄積パケット数の目標値を求めることと、時系列のゆらぎ時間のそれぞれに対応する蓄積パケット数の目標値を求めて、一定時間における蓄積パケット数の目標値の最大値を求めることは等価である。

。目標値の領域で最大値を求める方(後者)が、最大値を検索する処理を効率化できる場合もある。

[0043] また、パケットロスから蓄積パケット数の目標値を決める方法の例は以下のとおりである。紛失状態判定部233は、紛失検出部219から時々刻々送られたパケットロスの情報を用いて、フレームごと(例えば20ミリ秒ごと)に過去一定時間(例えば2秒間)における連続パケットロス数の最大値を求める。パケットロス情報の時系列から連続パケットロス数の最大値を求める方法は、パケットロス情報を一定時間蓄積して、その中で連続パケットロス数の最大値を検索するとよい。連続パケットロス数の最大値は、受信バッファ制御部231に送られる。受信バッファ制御部231は、得られた連続パケットロス数の最大値、あるいはそれに1を加えたフレーム数分を蓄積パケット数の目標値として、受信バッファ残量が目標値に近づくように制御する。蓄積パケット数の目標値はフレームごとに決定されるが、連続パケットロス数の最大値に変化がないときは、蓄積パケット数の目標値も同じ値が維持される。ゆらぎ時間と連続パケットロス数の両方から蓄積パケット数の目標値を決める方法としては、ゆらぎ時間から決まる蓄積パケット数の目標値と、連続パケットロス数から決まる蓄積パケット数の目標値の大きい方を、蓄積パケット数の目標値とする方法が考えられる。

[0044] 図20は、図18に示した音声信号受信装置200-2の変形例である。この機能構成は、双方向で音声信号をパケット通信網で通信する場合、パケットは双方向に送られることになるが、双方向の通信網の状態が同じと仮定している。そして、通信先から送られてきたパケットのゆらぎ時間や連続パケットロス数から、音声信号受信装置200-3とセットで備えられる音声信号送信装置の遅延量を決め、通信先に遅延量制御情報の設定依頼情報を送らない。この機能構成の場合は、音声信号送信装置100または100-2の端子118への入力はない。

[0045] 図21と図22は、連続パケットロス数のみから蓄積パケット数を決める例である。図18や図20との違いは、ゆらぎ検出部232が削除された点である。

図23と図24は、受信バッファ212の蓄積パケット数を、受信したパケット内に組み込まれた遅延量制御情報から決める場合の構成例である。この場合には、受信バッファ制御部231が削除される。例えば、受信バッファ212は、蓄積パケット数を、パケ

ット内に組み込まれた遅延量制御情報よりも1大きくする。また、図23の場合は、ゆらぎ情報が、受信状態を通信先に伝えるための情報である。図24の場合は、連続パケットロス数が受信状態を通信先に伝えるための情報である。

音声信号通信システム

図25に、一方向に音声パケットを送る場合のシステム構成例を示す。通信装置1000-1は音声信号送信装置100-2と記録部110から構成される。なお、音声信号送信装置は、100の音声信号送信装置でもよい。記録部110は、あらかじめ少なくとも遅延量制御情報を記録しており、音声信号送信装置100-2(または100)の端子117に遅延量制御情報を渡す。受信側の通信装置1000-2は、音声信号受信装置200-2から構成される。なお、音声信号受信装置は、200、200-4、200-6、あるいは200-7の音声信号受信装置でもよい。また、受信バッファに蓄積するパケット数の初期値として、例えば1を設定しておけばよい。このシステム構成のように完全な一方向の通信の場合には、音声信号受信装置200-2から送信装置への出力信号は不要である。したがって、音声信号受信装置200からバッファ残量復号部216と受信バッファ残量判定部218を削除した構成、音声信号受信装置200-2と200-4から遅延量制御部341を削除した構成、音声信号受信装置200-6から遅延量制御部341とゆらぎ検出部232を削除した構成、音声信号受信装置200-7から遅延量制御部341と紛失状態判定部233を削除した構成でもよい。

- [0046] 図26に、音声パケットを一方向で送り、遅延量制御情報の設定依頼情報を逆方向に送る場合のシステム構成例を示す。この場合、遅延量制御情報の設定依頼情報を送る通信網は、音声パケットの通信網と同じIP通信網13でもよいし、異なる通信網13-2でもよい。音声パケットの送信側の通信装置1000-3は、音声信号送信装置100-2と受信装置400から構成される。なお、音声信号送信装置は、100の音声信号送信装置でもよい。また、遅延量の初期値として、例えば1を設定しておけばよい。音声パケットの受信側の通信装置1000-4は音声信号受信装置200-2と送信装置300から構成される。なお、音声信号受信装置は、200、200-4、200-6、あるいは200-7の音声信号受信装置でもよい。また、蓄積パケット数の初期値として、例えば1を設定しておけばよい。

[0047] 図27に前記の送信装置300の機能構成例を、図28に前記の受信装置400の機能構成例を示す。送信装置300の端子320は、音声信号送信装置100-2の端子118に相当する。音声信号受信装置200-2からの情報は端子320を経由して送信部310に入力される。送信部310は、受け取った情報を通信網13(または13-2)の protocols にあわせたフォーマットに変更し、通信装置1000-3に送信する。ここで、音声信号受信装置からの情報とは、遅延量制御情報の設定依頼情報(遅延量制御情報の変更を依頼する情報、蓄積パケット数、バッファ残量、パケットのゆらぎ、あるいは連続パケットロス数など)である。受信装置400は、音声パケットの受信側の通信装置1000-4からの情報を受け取る。受け取った情報は受信部310に入力され、受信側の通信装置1000-4を制御するための情報は、音声信号送信装置の端子118へ渡される。また、送信側の通信装置1000-3内部の制御に使う情報は、遅延量制御部341に渡される。遅延量制御部341は、受け取った情報から遅延量制御情報を生成し、音声信号送信装置の端子117へ渡す。

[0048] 図29に、双方向で音声パケットを送る場合のシステム構成例を示す。通信装置1000-5は、音声信号送信装置100-2と音声信号受信装置200-2から構成される。なお、遅延量と受信バッファに蓄積する蓄積パケット数の初期値は、例えば1と設定しておけばよい。また、音声信号送信装置は、100の音声信号送信装置でもよい。音声信号受信装置は、200、200-3、200-4、200-5、200-6、あるいは200-7の音声信号受信装置でもよい。

以下に、上記のシステム構成での処理フローを説明するが、装置内部の細かい処理フローは図9および図10で示しているため省略する。例えば、音響信号を復号化する処理や紛失フレームの生成方法などは、実際の装置の内部では必ず行うことである。しかし、細かな処理は本発明の主要部ではないので、省略している。

[0049] 図30は、音声パケットの受信側から送信側への遅延量制御情報の設定依頼情報の送信を行わない場合の処理フローを示している。この処理フローは、図25のシステム構成の場合だけでなく、図26や図29のシステム構成の場合にも適用できる。送信側の通信装置1000-1に音響信号が入力されると入力バッファ111に蓄積される(S101)。蓄積された音響信号は、フレームごとに分割され、音声符号化される(S10

2)。次に音響信号対応データと遅延量制御情報とを含む音声パケットを構成し、受信側に送信する(S103)。受信側の通信装置1000-2では、音声パケットを受信し(S201)、受信バッファ212に蓄積する(S202)。パケットロスが発生した場合には(S203)、紛失処理制御部217で遅延量制御情報を利用して音響信号対応データを検索し(S204)、紛失フレームを生成する(S205)。最後にフレームの順番に音響信号を並べて出力する(S206)。このシステムでは、送信側の通信装置1000-1が複数ある場合に、それぞれの通信装置1000-1が自由に遅延量を設定することができる。また、図29の通信装置1000-5の場合にも、通信装置ごとに自由に遅延量を設定、変更することができる。したがって、遅延量についてあらかじめ共通のルールを決めておく必要がないし、変更も可能であるという効果がある。

[0050] 図31、図32、図33、および図34は、音声パケットの受信側から送信側への遅延量制御情報の設定依頼情報の送信を行う場合の処理フローを示している。これらの処理フローは図26と図29のシステム構成の場合に適用できる。

図31の処理フローは、少なくとも、音声パケットの送信側の通信装置1000-3(または1000-5)が音声信号送信装置100-2を備えており、音声パケットの受信側の通信装置1000-4(または1000-5)が音声信号受信装置200-2または200-4を備えている場合に適用できる。この処理フローの特徴は、受信側で、ゆらぎ時間または連続パケットロス数から遅延量制御情報の設定依頼情報を決め、送信側で、遅延量制御情報の設定依頼情報から遅延量制御情報を求めることである。

[0051] 音声パケットの送信側の通信装置1000-3に、音響信号が入力され(S101)、入力バッファ111に蓄積される(S102)。音響信号対応データと遅延量制御情報とが組み込まれた音声パケットが構成され、受信側に送信される(S103)。音声パケットの受信側の通信装置1000-4では、音声パケットを受信し(S201)、受信バッファ212に蓄積する(S202)。受信側の通信装置1000-4のゆらぎ検出部232でゆらぎ時間を検出(または、紛失状態判定部233で連続パケットロス数を検出)する(S207)。受信バッファ制御部231は、蓄積パケット数の目標値を、ゆらぎ時間または連続パケットロス数以上に設定する(S208)。蓄積パケット数の目標値は、音声パケットの受信側の通信装置1000-4の送信装置300(または、通信装置1000-5の音声信号

送信装置100-2)に遅延量制御情報の設定依頼情報として渡される。そして、送信装置300(または音声信号送信装置100-2)は、音声パケットの送信側の通信装置1000-3(または通信装置1000-5)に遅延量制御情報の設定依頼情報として送信する(S209)。音声パケットの送信側の通信装置1000-3では、遅延量制御情報の設定依頼情報を受信する(S104)。遅延量制御部341で遅延量制御情報の値を遅延量制御情報の設定依頼情報すなわち蓄積パケット数の目標値以下とし、音声信号送信装置100-2に渡す。音声信号送信装置100-2では、シフトバッファ116に遅延量制御情報が入力され、遅延量制御情報が設定される(S105)。音声パケットの受信側の通信装置1000-4では、パケットロス発生の確認、紛失フレームの生成などを行い(S203、S204、S205)、音声信号を出力する(S206)。このような処理によって、図30で示した処理フローの効果の他に、時々刻々と変化する通信網の状態に追従した最適な蓄積パケット数と遅延量を設定することができるという効果がある。

[0052] 図32の処理フローは、少なくとも、音声パケットの送信側の通信装置1000-3(または1000-5)が音声信号送信装置100-2を備えており、音声パケットの受信側の通信装置1000-4(または1000-5)が音声信号受信装置200-2、200-4、200-6、または200-7を備えている場合に適用できる。この処理の特徴は、次のとおりである。音声パケットの受信側で、ゆらぎ時間または連続パケットロス数を検出し、その結果を送信側に送る。音声パケットの送信側では、ゆらぎ時間または連続パケットロス数から遅延量制御情報を設定する。音声パケットの送信側から、音声信号と一緒に遅延量制御情報が受信側に送られる。音声パケットの受信側では、受信した遅延量制御情報から蓄積パケット数を設定する。

[0053] ステップS101からS201までは、図31の処理フローと同じである。受信バッファ212は、受信した音声パケットから、遅延量制御情報を取り出し、蓄積パケット数を設定する(S208)。この設定にしたがって、音声パケットは受信バッファに蓄積される(S202)。ゆらぎ検出部232または紛失状態判定部233は、ゆらぎ時間または連続パケットロス数を検出する(S207)。検出された情報は、音声パケットの受信側の通信装置1000-4内の送信装置300(または、通信装置1000-5内の音声信号送信装置

100-2)から、音声パケットの送信側の通信装置1000-3(または、通信装置1000-5)に送信される(S209)。音声パケットの送信側の通信装置1000-3(または、通信装置1000-5)では、情報を受信する(S106)。遅延量制御部341で、遅延量制御情報の値を、ゆらぎ時間に対応するフレーム数または連続パケットロス数よりも大きい値とし、音声信号送信装置100-2に渡す。音声信号送信装置100-2では、シフトバッファ116に遅延量制御情報が入力され、遅延量制御情報が設定される(S105)。ステップS203からS206は、図31と同じである。本処理フローの特徴の説明では、蓄積パケット数の設定が最後だったが、図32の処理フローでは、ゆらぎの検出や遅延量制御情報の設定よりも前となっている。これは、本処理が繰り返し行われることによるもので、蓄積パケット数は、1回前の繰り返し処理で求められた遅延量制御情報から決まる。つまり、蓄積パケット数の設定が最後である。なお、1回目の繰り返し処理の時には、遅延量制御情報の初期値として、例えば1を設定しておけばよい。この処理フローの効果は図31の場合と同じである。

[0054] 図33の処理フローは、少なくとも、音声パケットの送信側の通信装置1000-3(または1000-5)が音声信号送信装置100-2を備えており、音声パケットの受信側の通信装置1000-4(または1000-5)が音声信号受信装置200-2または200-4を備えている場合に適用できる。この処理の特徴は、ゆらぎ時間または連続パケットロス数から蓄積パケット数の目標値と遅延量制御情報とを同時に求めることである。

ステップS101からS207までは図31と同じである。検出されたゆらぎまたは連続パケットロス数から、受信バッファ制御部231で蓄積パケット数の目標値と遅延量制御情報とを求める(S210)。このとき、

ゆらぎ時間に相当するフレーム数(端数切り上げ) \leq 遅延量 \leq 蓄積パケット数

または

連続パケットロス数 \leq 遅延量 \leq 蓄積パケット数

の関係を満足するように求める。受信バッファ212は、求められた蓄積パケット数の目標値にしたがってパケットを蓄積するように設定される(S208)。求められた遅延量制御情報は、音声パケットの受信側の通信装置1000-4の送信装置300(または、通信装置1000-5の音声信号送信装置100-2)から、音声パケットの送信側の通

信装置1000-3(または、通信装置1000-5)に、遅延量制御情報の変更依頼として送信される(S211)。この情報を、音声パケットの送信側の通信装置1000-3(または、通信装置1000-5)が受信し(S107)、遅延量制御部341で遅延量制御情報とする。音声信号送信装置100-2では、シフトバッファ116に遅延量制御情報が入力され、遅延量制御情報が設定される(S105)。ステップS203からS206は、図31と同じである。この処理フローの効果は、図31、図32の場合と同じである。

[0055] 図34の処理フローは、少なくとも、音声パケットの送信側の通信装置1000-3(または1000-5)が音声信号送信装置100を備えており、音声パケットの受信側の通信装置1000-4(または1000-5)が音声信号受信装置200を備えている場合に適用できる。この処理の特徴は、バッファ残量を遅延量制御情報の設定依頼情報として送信し、遅延量制御情報を求めることである。

ステップS101からステップS202までは図31と同じである。音声パケットの受信側の通信装置1000-4(または、通信装置1000-5)の受信バッファ残量判定部218は、受信バッファの残量を判定する(S213)。判定された情報は、音声パケットの受信側の通信装置1000-4の送信装置300(または、通信装置1000-5の音声信号送信装置100-2)から、音声パケットの送信側の通信装置1000-3(または、通信装置1000-5)に送信される(S214)。音声パケットの送信側の通信装置1000-3(または、通信装置1000-5)では、バッファ残量の情報を受信する(S108)。遅延量制御部341(または、バッファ残量復号部216)は、バッファ残量より以下となる遅延量制御情報を求める。音声信号送信装置100-2では、シフトバッファ116に遅延量制御情報が入力され、遅延量制御情報が設定される(S105)。ステップS203からS206は、図31と同じである。この処理フローの効果は、図31、図32、図33の場合と同じである。

[0056] 図35は、双方向で音声パケットをやり取りする場合に、通信網の状態がパケット通信に与える影響は、双方向とも同じであると仮定した処理フローである。この処理フローは、図29のシステム構成の場合に適用できる。この処理の特徴は、受信した音声パケットのゆらぎや連続パケットロス数から、受信するパケットの蓄積パケット数と送信する音声パケットの遅延量制御情報を決めることである。ステップS101からステップ

S208までは図33と同じである。求められた遅延量制御情報は、音声信号送信装置100-2のシフトバッファ116に入力され、遅延量制御情報が設定される(S212)。この処理フローの効果は、図31、図32、図33の効果に加え、通信装置を制御するための手順が少なくなることである。

- [0057] 図17、18、20～24、27、28に示した装置はそれぞれコンピュータにより機能させてもよい。図30から35に示した処理手順の各過程は、プログラムによってコンピュータに実行させてもよい。また、プログラムは、コンピュータにCD-ROM、磁気ディスク装置、半導体記憶装置などの記録媒体からインストール、あるいは通信回線を介してダウンロードすればよい。

請求の範囲

- [1] 1つの通信装置は少なくとも送信部を備え、対する他の1つ以上の通信装置は少なくとも受信部を備え、これらの通信装置の間でパケット通信を行う場合に、
- 送信部で、
- 音声・音楽などの音響信号をフレームと呼ばれる一定時間ごとに区切ってフレーム音響信号を生成する過程と、
- 前記フレーム音響信号から前記フレーム音響信号に対応するデータ(以下、「音響信号対応データ」という。)を生成する過程と、
- 前記フレーム音響信号と前記音響信号対応データをパケットに格納して送信する過程と、
- 受信部で、
- 受信したパケットを受信バッファに蓄える過程と、
- 取り出すべきフレーム番号を指定する過程と、
- 前記取り出すべきフレーム番号に対応するフレーム音響信号を含むパケットが受信バッファに格納されているか、格納されていないかを判定する紛失検出過程と、
- 前記紛失検出過程で、前記取り出すべきフレーム番号に対応するフレーム音響信号を含むパケットが受信バッファに格納されていると判定された場合には、受信バッファに格納されている当該パケットからフレーム音響信号を取り出してフレーム出力音響信号とする音響パケット復号過程と、
- 前記紛失検出過程で、前記取り出すべきフレーム番号に対応するフレーム音響信号を含むパケットが受信バッファに格納されていないと判定された場合(以下、「パケットロスが発生した場合」)には、当該フレーム(以下、「紛失フレーム」)の音響信号対応データを受信バッファに格納されたパケットから取り出し、該音響信号対応データを用いてフレーム出力音響信号を生成する紛失処理過程と、
- 前記音響パケット復号過程または前記紛失処理過程から出力されたフレーム出力音響信号を連結して出力する過程と、
- を有する音響信号パケット通信方法において、
- 前記送信部で、フレーム音響信号と同一のパケットに、遅延量制御情報で指定す

る値だけフレーム番号が異なるフレームの音響信号対応データと、遅延量制御情報とを組み込んで送信する過程と、

前記受信部で、パケットロスが発生した場合に、紛失フレームと同じフレーム番号の音響信号対応データを、パケットに組み込まれた前記遅延量制御情報を用いて前記受信バッファ内のパケットから求める過程と、

を有することを特徴とする音響信号パケット通信方法。

[2] 請求項1記載の音響信号パケット通信方法において、

1つの通信装置に前記送信部と前記受信部の両方を備え、対する他の1つ以上の通信装置も前記送信部と前記受信部の両方を備え、これらの通信装置の間でパケット通信を行う場合に、

受信部で、

受信されるパケットのゆらぎ状態を判定する第一の判定過程と、受信されるパケットの紛失状態を判定する第二の判定過程の両方あるいはいずれか一方と、

前記判定過程における判定結果を用いて、受信バッファに蓄積すべきパケットの数(以下、「蓄積パケット数の目標値」という。)を決定する過程

を有し、

前記受信部と同一の通信装置内の送信部で、

前記遅延量制御情報を前記蓄積パケット数の目標値以下に設定する過程

を有する

ことを特徴とする音響信号パケット通信方法。

[3] 請求項1記載の音響信号パケット通信方法において、

1つの通信装置に前記送信部と前記受信部の両方を備え、対する他の1つ以上の通信装置も前記送信部と前記受信部の両方を備え、これらの通信装置の間でパケット通信を行う場合に、

受信部で、

受信されるパケットのゆらぎ状態を判定する第一の判定過程と、受信されるパケットの紛失状態を判定する第二の判定過程の両方あるいはいずれか一方と、

前記判定過程における判定結果を用いて、受信バッファに蓄積すべきパケットの数

(以下、「蓄積パケット数の目標値」という。)を決定する過程と、
前記蓄積パケット数の目標値を、同一の通信装置内の送信部に送る過程と
を有し、
前記受信部と同一の通信装置内の送信部で、
前記受信部から送られた蓄積パケット数の目標値を、通信先の送信部において設定される遅延量制御情報を指定する情報として、パケットに格納して送信する過程と
を有する
ことを特徴とする音響信号パケット通信方法。

- [4] 請求項1記載の音響信号パケット通信方法において、
1つの通信装置に前記送信部と前記受信部の両方を備え、対する他の1つ以上の通信装置も前記送信部と前記受信部の両方を備え、これらの通信装置の間でパケット通信を行う場合に、
受信部で、
受信バッファに蓄積されているパケットの数(以下、「バッファ残量」という。)を測定する過程と、
前記バッファ残量を、同一の通信装置内の送信部に送る過程と
を有し、
前記受信部と同一の通信装置内の送信部で、
前記受信部から送られたバッファ残量を、通信先の送信部において設定される遅延量制御情報を指定する情報として、パケットに格納して送信する過程と
を有する
ことを特徴とする音響信号パケット通信方法。

- [5] 1つの通信装置は少なくとも送信部を備え、対する他の1つ以上の通信装置は少なくとも受信部を備え、これらの通信装置の間でパケット通信を行う場合に、
送信部で、
音声・音楽などの音響信号をフレームと呼ばれる一定時間ごとに区切ってフレーム音響信号を生成する過程と、
前記フレーム音響信号から前記フレーム音響信号に対応するデータ(以下、「音響

信号対応データ」という。)を生成する過程と、

前記フレーム音響信号と前記音響信号対応データをパケットに格納して送信する過程と、

を有する音響信号パケット送信方法において、

前記送信部で、フレーム音響信号と同一のパケットに、遅延量制御情報で指定する値だけフレーム番号が異なるフレームの音響信号対応データと、遅延量制御情報とを組み込んで送信する過程

を有することを特徴とする音響信号パケット送信方法。

[6] 請求項5記載の音響信号パケット送信方法において、

1つの通信装置に前記送信部と前記受信部の両方を備え、対する他の1つ以上の通信装置も前記送信部と前記受信部の両方を備え、これらの通信装置の間でパケット通信を行う場合に、

送信部で、

当該送信部と同一の通信装置内の受信部で決定した当該受信部に蓄積すべきパケットの数以下に前記遅延量制御情報を設定する過程を有する

ことを特徴とする音響信号パケット送信方法。

[7] 請求項5記載の音響信号パケット送信方法において、

1つの通信装置に前記送信部と前記受信部の両方を備え、対する他の1つ以上の通信装置も前記送信部と前記受信部の両方を備え、これらの通信装置の間でパケット通信を行う場合に、

送信部で、

当該送信部と同一の通信装置内の受信部で決定した当該受信部に蓄積すべきパケットの数を、通信先の送信部において設定される遅延量制御情報を指定する情報として、パケットに格納して送信する過程と

を有する

ことを特徴とする音響信号パケット送信方法。

[8] 請求項5記載の音響信号パケット送信方法において、

1つの通信装置に前記送信部と前記受信部の両方を備え、対する他の1つ以上の

通信装置も前記送信部と前記受信部の両方を備え、これらの通信装置の間でパケット通信を行う場合に、

送信部で、

当該送信部と同一の通信装置内の受信部で測定した受信バッファに蓄積されているパケットの数を、通信先の送信部において設定される遅延量制御情報の設定依頼情報として、パケットに格納して送信する過程を有する

ことを特徴とする音響信号パケット送信方法。

[9] 1つの通信装置は少なくとも送信部を備え、対する他の1つ以上の通信装置は少なくとも受信部を備え、これらの通信装置の間でパケット通信を行う場合に、

受信部で、

受信したパケットを受信バッファに蓄える過程と、

取り出すべきフレーム番号を指定する過程と、

前記取り出すべきフレーム番号に対応するフレーム音響信号を含むパケットが受信バッファに格納されているか、格納されていないかを判定する紛失検出過程と、

前記紛失検出過程で、前記取り出すべきフレーム番号に対応するフレーム音響信号を含むパケットが受信バッファに格納されていると判定された場合には、受信バッファに格納されている当該パケットからフレーム音響信号を取り出してフレーム出力音響信号とする音響パケット復号過程と、

前記紛失検出過程で、前記取り出すべきフレーム番号に対応するフレーム音響信号を含むパケットが受信バッファに格納されていないと判定された場合(以下、「パケットロスが発生した場合」)には、当該フレーム(以下、「紛失フレーム」)の音響信号に対応するデータ(以下、「音響信号対応データ」という。)を受信バッファに格納されたパケットから取り出し、該音響信号対応データを用いてフレーム出力音響信号を生成する紛失処理過程と、

前記音響パケット復号過程または前記紛失処理過程から出力されたフレーム出力音響信号を連結して出力する過程と、

を有する音響信号パケット受信方法において、

前記受信部で、パケットロスが発生した場合に、紛失フレームと同じフレーム番号の

音響信号対応データを、パケットに組み込まれた遅延量制御情報を用いて前記受信バッファ内のパケットから求める過程

を有することを特徴とする音響信号パケット受信方法。

[10] 請求項9記載の音響信号パケット受信方法において、

1つの通信装置に前記送信部と前記受信部の両方を備え、対する他の1つ以上の通信装置も前記送信部と前記受信部の両方を備え、これらの通信装置の間でパケット通信を行う場合に、

受信部で、

受信されるパケットのゆらぎ状態を判定する第一の判定過程と、受信されるパケットの紛失状態を判定する第二の判定過程の両方あるいはいずれか一方と、

前記判定過程における判定結果を用いて、受信バッファに蓄積すべきパケットの数を決定する過程と、

前記受信バッファに蓄積すべきパケットの数を、同一の通信装置内の送信部に送る過程と、

を有することを特徴とする音響信号パケット受信方法。

[11] 請求項9記載の音響信号パケット受信方法において、

1つの通信装置に前記送信部と前記受信部の両方を備え、対する他の1つ以上の通信装置も前記送信部と前記受信部の両方を備え、これらの通信装置の間でパケット通信を行う場合に、

受信部で、

受信バッファに蓄積されているパケットの数(以下、「バッファ残量」という。)を測定する過程と、

前記バッファ残量を、同一の通信装置内の送信部に送る過程と

を有する

ことを特徴とする音響信号パケット受信方法。

[12] 1つの通信装置は少なくとも送信部を備え、対する他の1つ以上の通信装置は少なくとも受信部を備え、これらの通信装置の間でパケット通信を行う場合に、

送信部に、

音声・音楽などの音響信号をフレームと呼ばれる一定時間ごとに区切ってフレーム音響信号を生成する手段と、

前記フレーム音響信号から前記フレーム音響信号に対応するデータ(以下、「音響信号対応データ」という。)を生成する手段と、

前記フレーム音響信号と前記音響信号対応データをパケットに格納して送信する手段とを有し、

受信部に、

受信したパケットを受信バッファに蓄える手段と、

取り出すべきフレーム番号を指定する手段と、

前記取り出すべきフレーム番号に対応するフレーム音響信号を含むパケットが受信バッファに格納されているか、格納されていないかを判定する紛失検出手段と、

前記紛失検出手段で、前記取り出すべきフレーム番号に対応するフレーム音響信号を含むパケットが受信バッファに格納されていると判定された場合には、受信バッファに格納されている当該パケットからフレーム音響信号を取り出してフレーム出力音響信号とする音響パケット復号手段と、

前記紛失検出手段で、前記取り出すべきフレーム番号に対応するフレーム音響信号を含むパケットが受信バッファに格納されていないと判定された場合(以下、「パケットロスが発生した場合」)には、当該フレーム(以下、「紛失フレーム」)の音響信号対応データを受信バッファに格納されたパケットから取り出し、該音響信号対応データを用いてフレーム出力音響信号を生成する紛失処理手段と、

前記音響パケット復号手段または前記紛失処理手段から出力されたフレーム出力音響信号を連結して出力する手段とを有する

音響信号パケット通信システムにおいて、

前記送信部に、フレーム音響信号と同一のパケットに、遅延量制御情報で指定する値だけフレーム番号が異なるフレームの音響信号対応データと、遅延量制御情報とを組み込んで送信する手段を有し、

前記受信部に、パケットロスが発生した場合に、紛失フレームと同じフレーム番号の音響信号対応データを、パケットに組み込まれた前記遅延量制御情報を用いて前記

受信バッファ内のパケットから求める手段を有する

ことを特徴とする音響信号パケット通信システム。

- [13] 音声・音楽などの音響信号をフレームと呼ばれる一定時間ごとに区切ってフレーム音響信号を生成する手段と、

前記フレーム音響信号から前記フレーム音響信号に対応するデータ(以下、「音響信号対応データ」という。)を生成する手段と、

前記フレーム音響信号と前記音響信号対応データをパケットに格納して送信する手段とを有する送信部と、

受信したパケットを受信バッファに蓄える手段と、

取り出すべきフレーム番号を指定する手段と、

前記取り出すべきフレーム番号に対応するフレーム音響信号を含むパケットが受信バッファに格納されているか、格納されていないかを判定する紛失検出手段と、

前記紛失検出手段で、前記取り出すべきフレーム番号に対応するフレーム音響信号を含むパケットが受信バッファに格納されていると判定された場合には、受信バッファに格納されている当該パケットからフレーム音響信号を取り出してフレーム出力音響信号とする音響パケット復号手段と、

前記紛失検出手段で、前記取り出すべきフレーム番号に対応するフレーム音響信号を含むパケットが受信バッファに格納されていないと判定された場合(以下、「パケットロスが発生した場合」)には、当該フレーム(以下、「紛失フレーム」)の音響信号対応データを受信バッファに格納されたパケットから取り出し、該音響信号対応データを用いてフレーム出力音響信号を生成する紛失処理手段と、

前記音響パケット復号手段または前記紛失処理手段から出力されたフレーム出力音響信号を連結して出力する手段とを有する受信部と、

を備える音響信号パケット通信装置において、

前記送信部に、フレーム音響信号と同一のパケットに、遅延量制御情報で指定する値だけフレーム番号が異なるフレームの音響信号対応データと、遅延量制御情報とを組み込んで送信する手段を有し、

前記受信部に、パケットロスが発生した場合に、紛失フレームと同じフレーム番号の

音響信号対応データを、パケットに組み込まれた前記遅延量制御情報を用いて前記受信バッファ内のパケットから求める手段を有する

ことを特徴とする音響信号パケット通信装置。

[14] 請求項13記載の音響信号パケット通信装置において、

前記受信部に、受信されるパケットのゆらぎ状態を判定する第一の判定手段と、受信されるパケットの紛失状態を判定する第二の判定手段の両方あるいはいずれか一方と、

前記判定手段での判定結果を用いて、受信バッファに蓄積すべきパケットの数(以下、「蓄積パケット数の目標値」という。)を決定する手段とを有し、

前記送信部に、前記遅延量制御情報を前記蓄積パケット数の目標値以下に設定する手段を有する

ことを特徴とする音響信号パケット通信装置。

[15] 請求項13記載の音響信号パケット通信装置において、

前記受信部に、受信されるパケットのゆらぎ状態を判定する第一の判定手段と、受信されるパケットの紛失状態を判定する第二の判定手段の両方あるいはいずれか一方と、

前記判定手段での判定結果を用いて、受信バッファに蓄積すべきパケットの数(以下、「蓄積パケット数の目標値」という。)を決定する手段とを有し、

前記送信部に、前記蓄積パケット数の目標値を、通信先の送信部において設定される遅延量制御情報を指定する情報として、パケットに格納して送信する手段を有する

ことを特徴とする音響信号パケット通信装置。

[16] 請求項13記載の音響信号パケット通信装置において、

前記受信部に、受信バッファに蓄積されているパケットの数(以下、「バッファ残量」という。)を測定する手段を有し、

前記送信部に、前記バッファ残量を、通信先の送信部において設定される遅延量制御情報を指定する情報として、前記フレーム音響信号と同一のパケットに組み込んで送信する手段を有する

ことを特徴とする音響信号パケット通信装置。

- [17] 請求項1記載の音響信号パケット通信方法の各過程をコンピュータにより実行させるための音響信号パケット通信プログラム。
- [18] 請求項5記載の音響信号パケット送信方法の各過程をコンピュータにより実行させるための音響信号パケット送信プログラム。
- [19] 請求項9記載の音響信号パケット受信方法の各過程をコンピュータにより実行させるための音響信号パケット受信プログラム。

[図1]

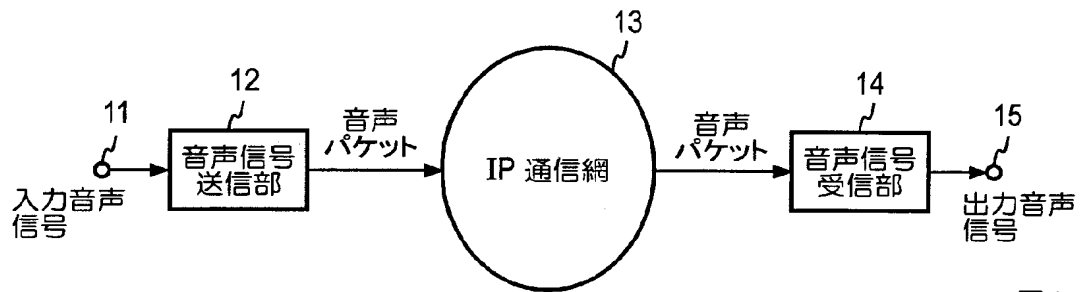


図1

[図2]

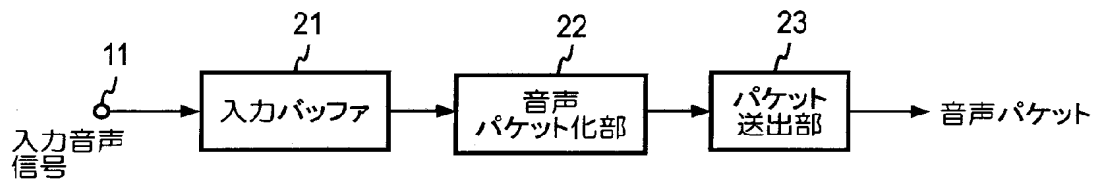


図2

[図3]

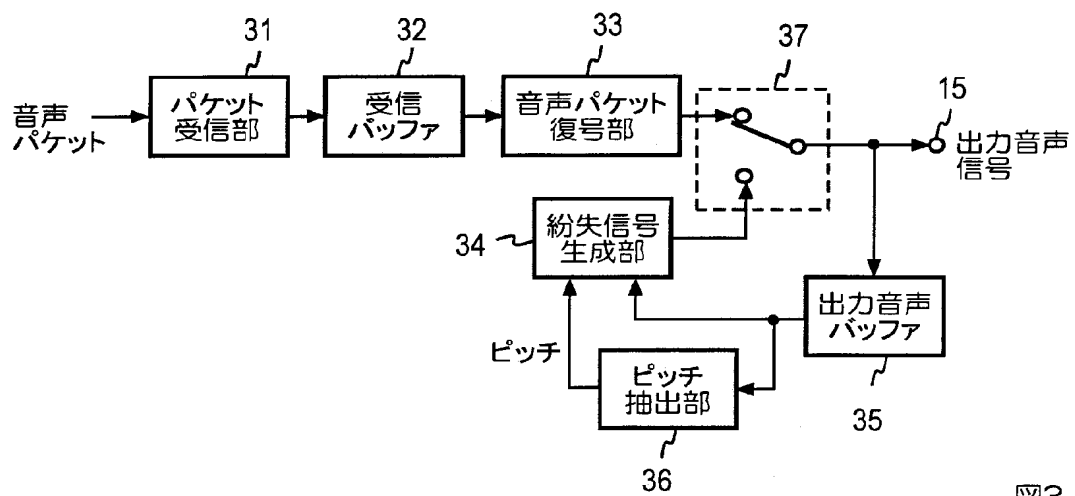


図3

[図4]

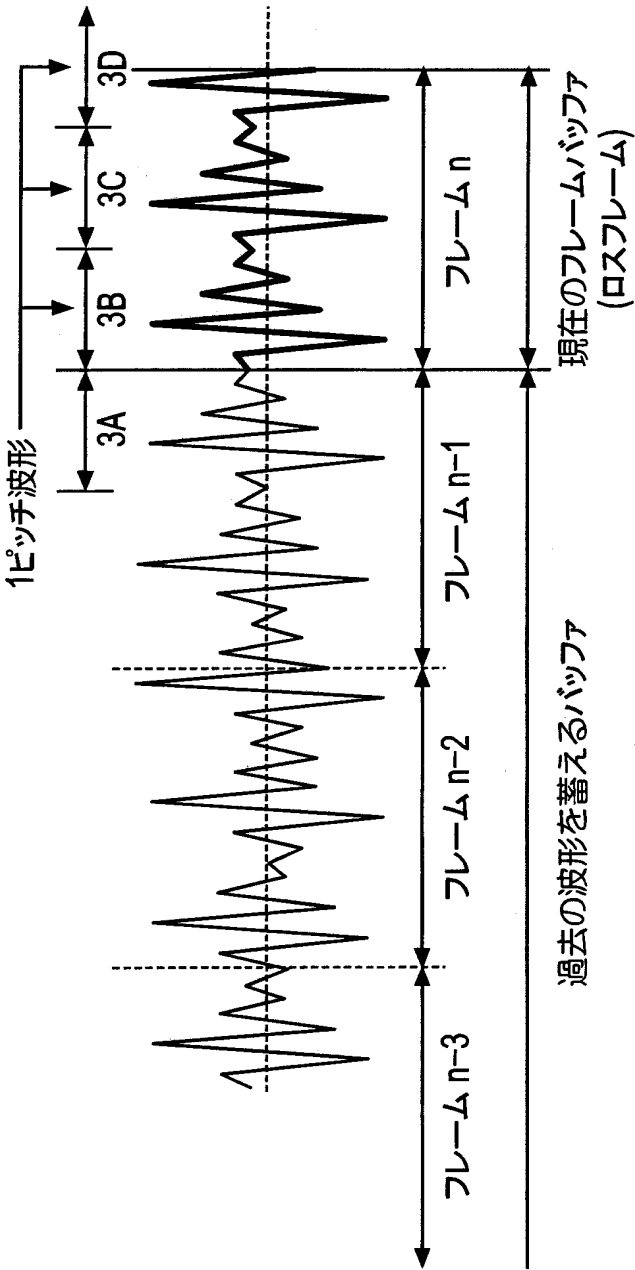
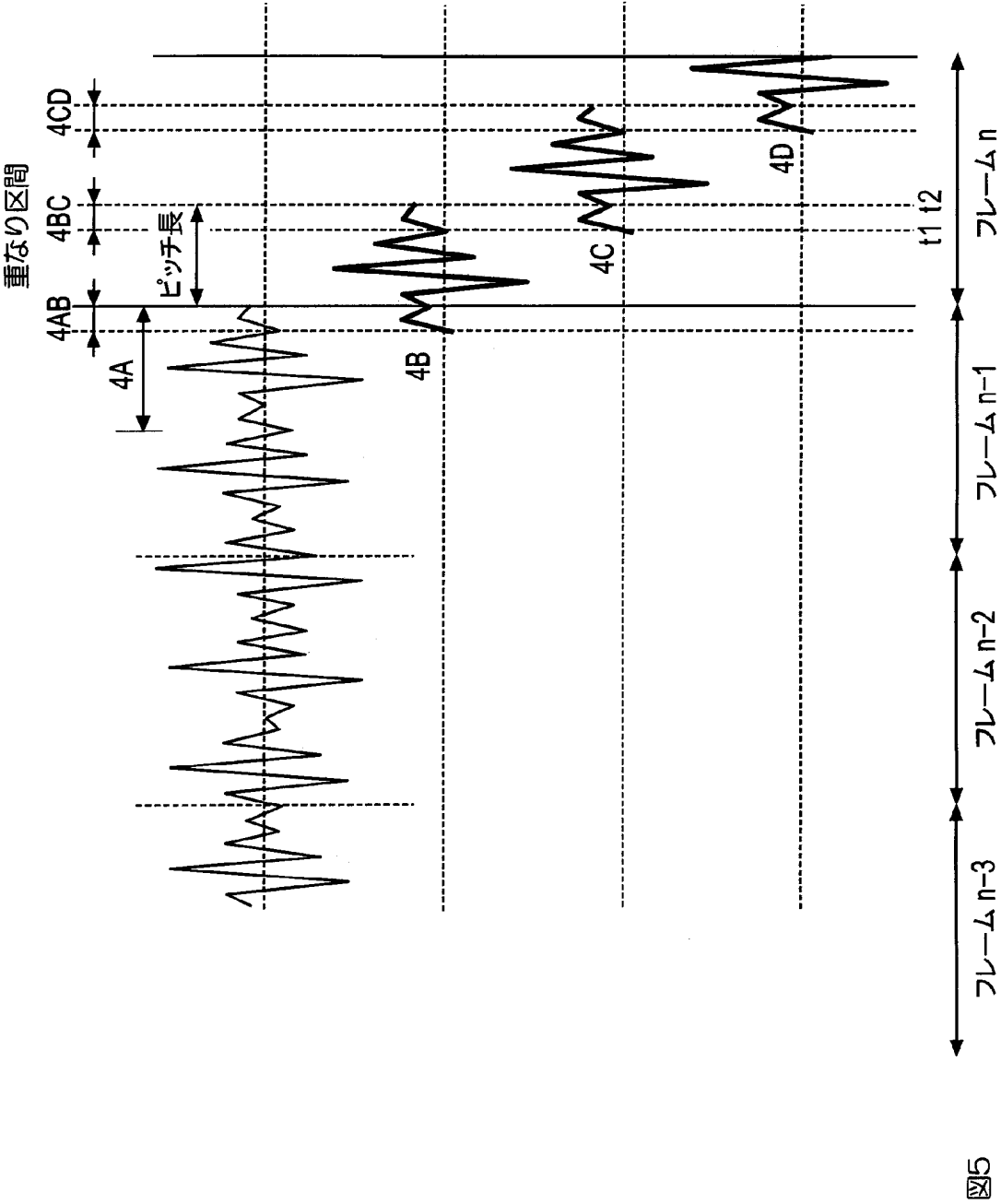


図4

[図5]



[図6]

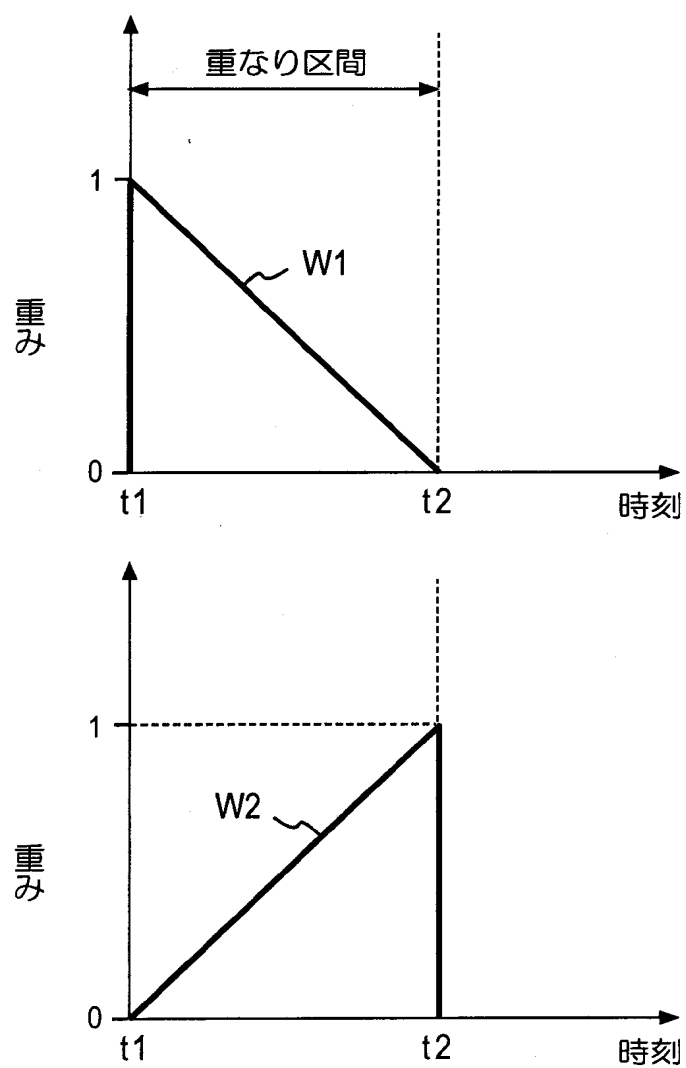


図6

[図7]

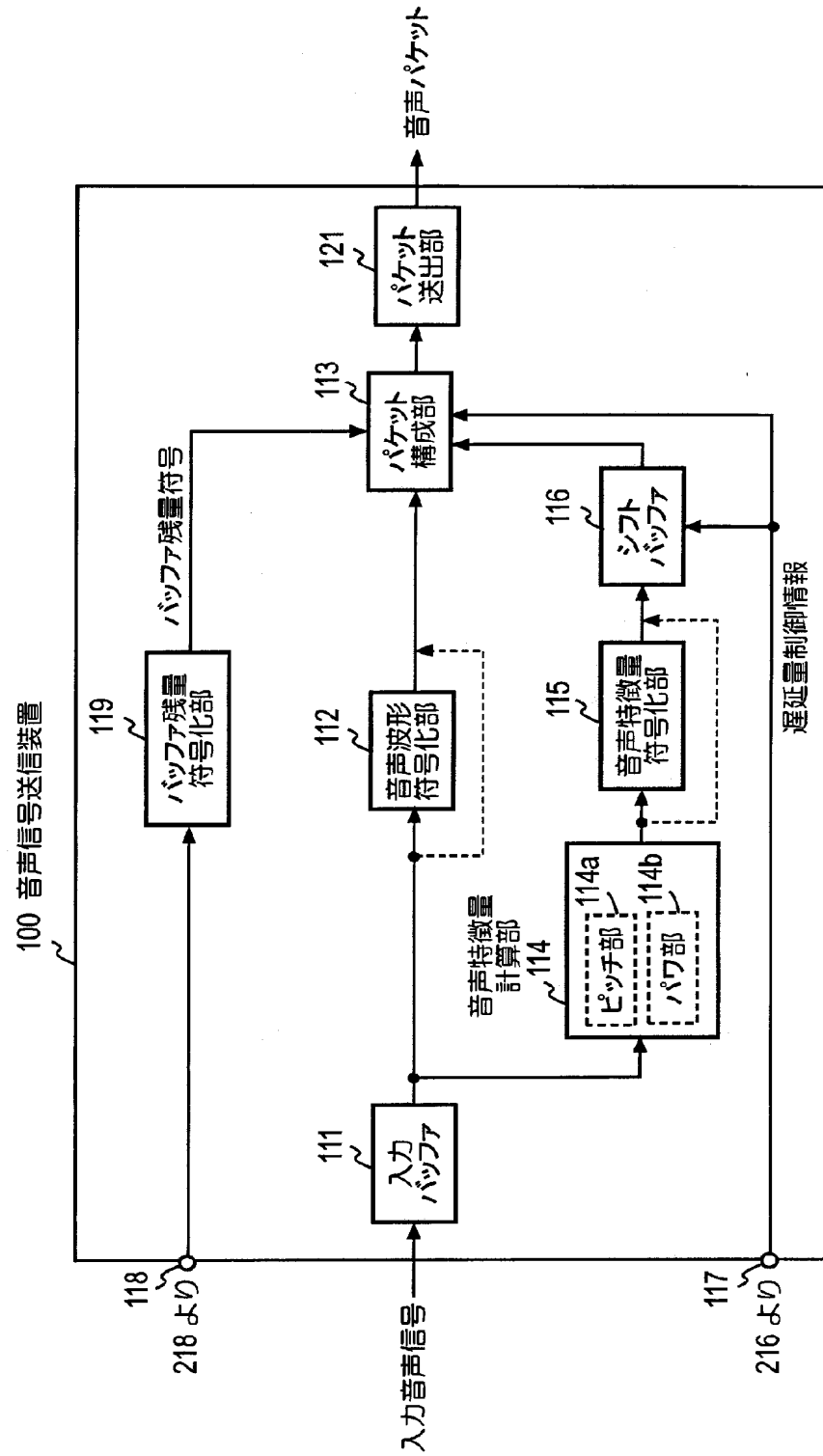
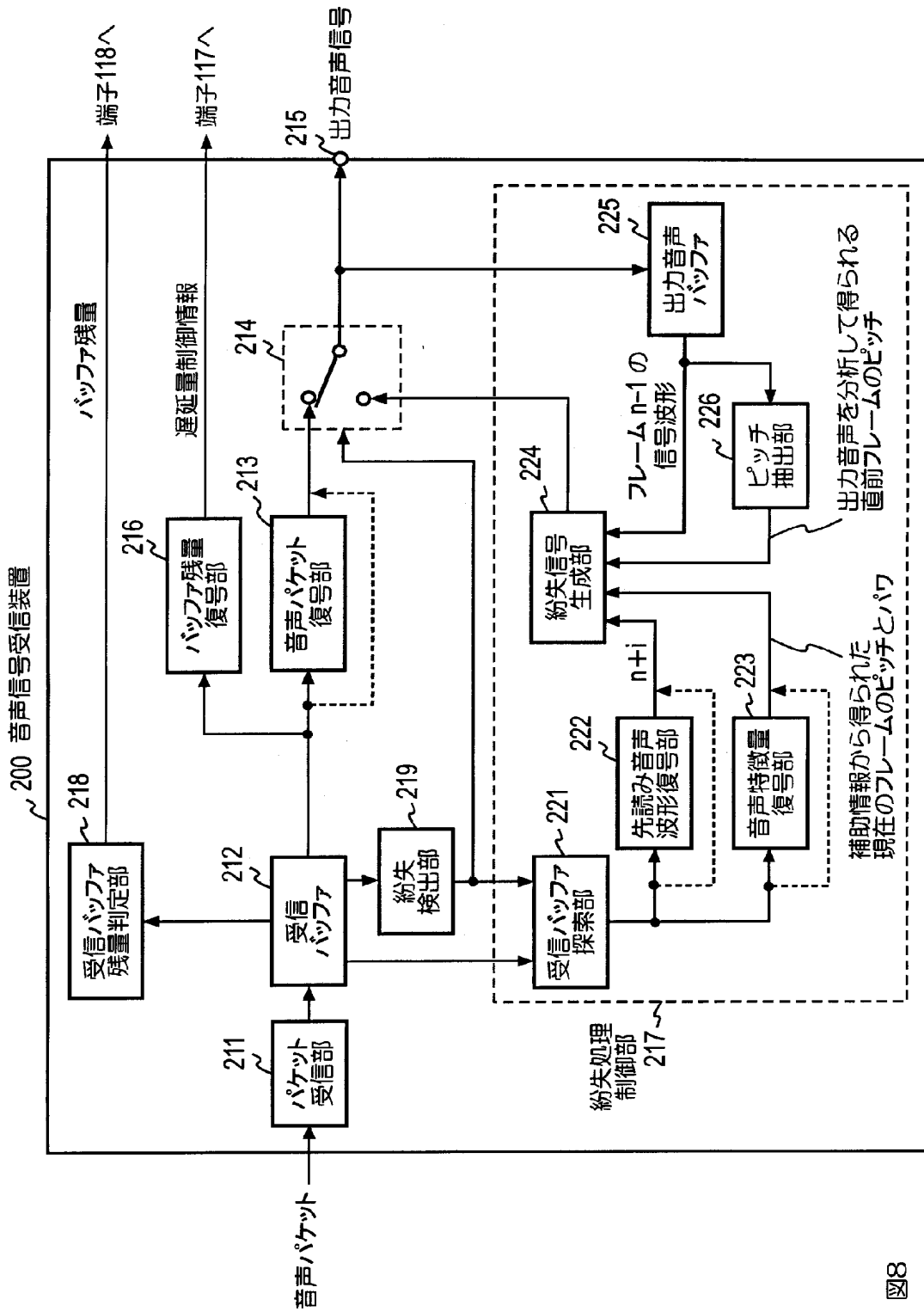


図7

[图8]



[図9]

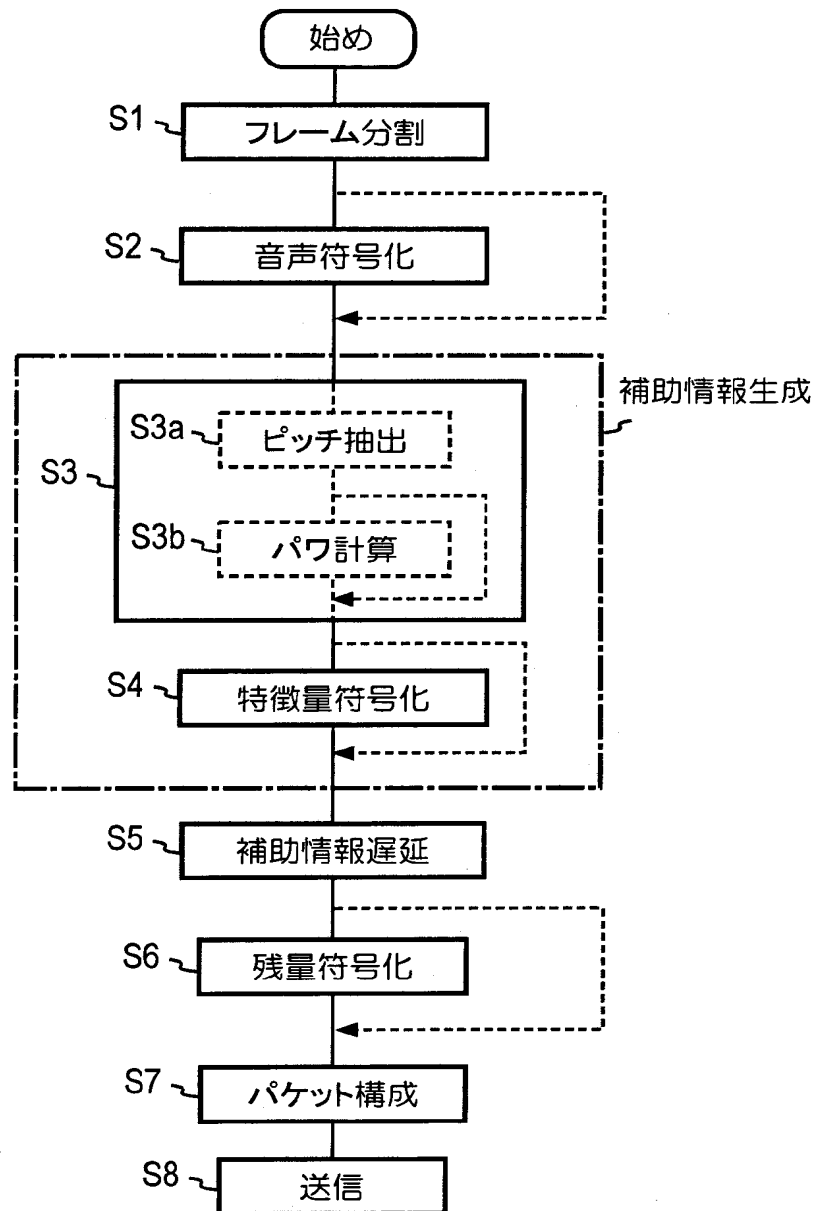


図9

[図10]

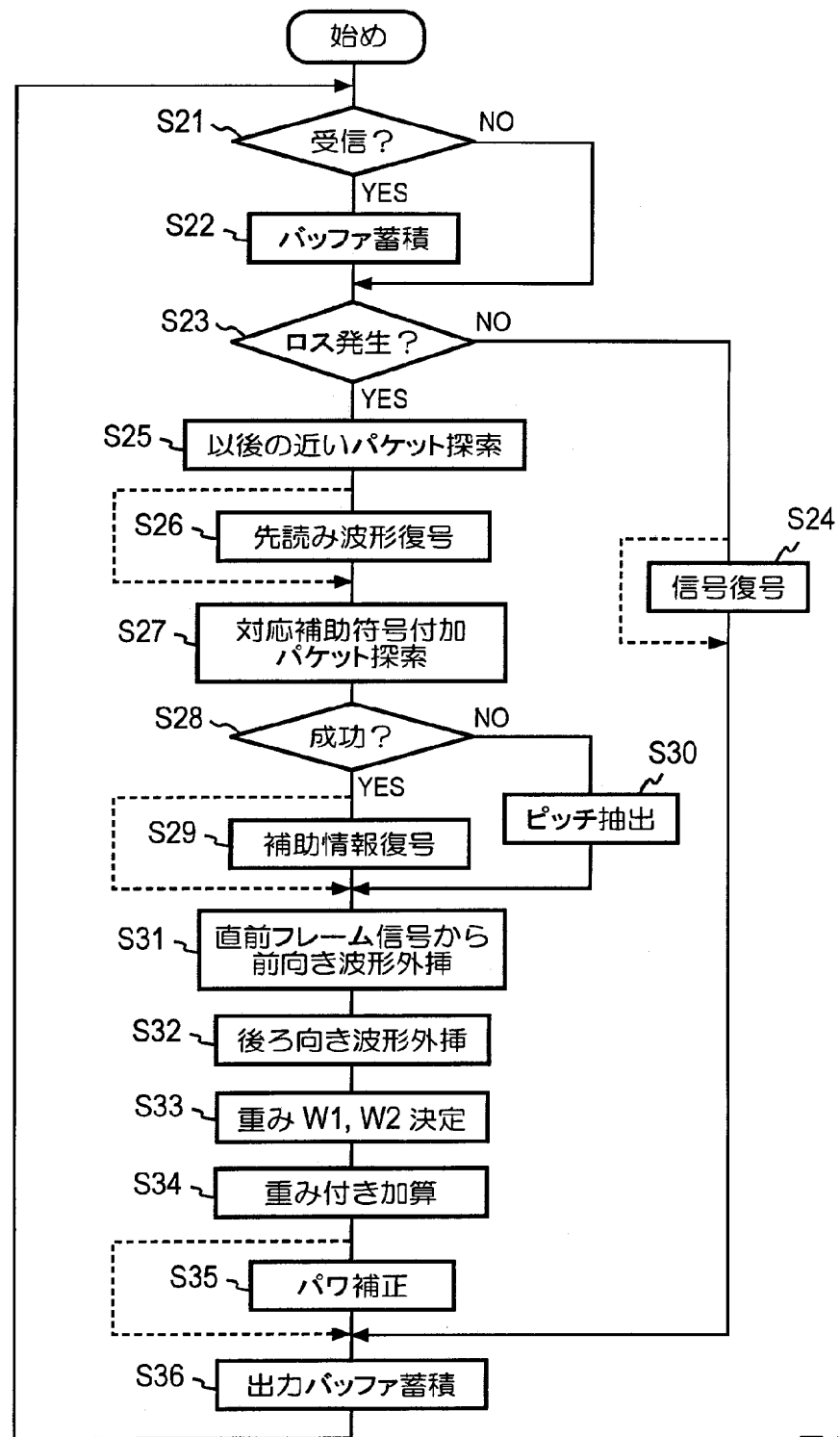


図10

[図11]

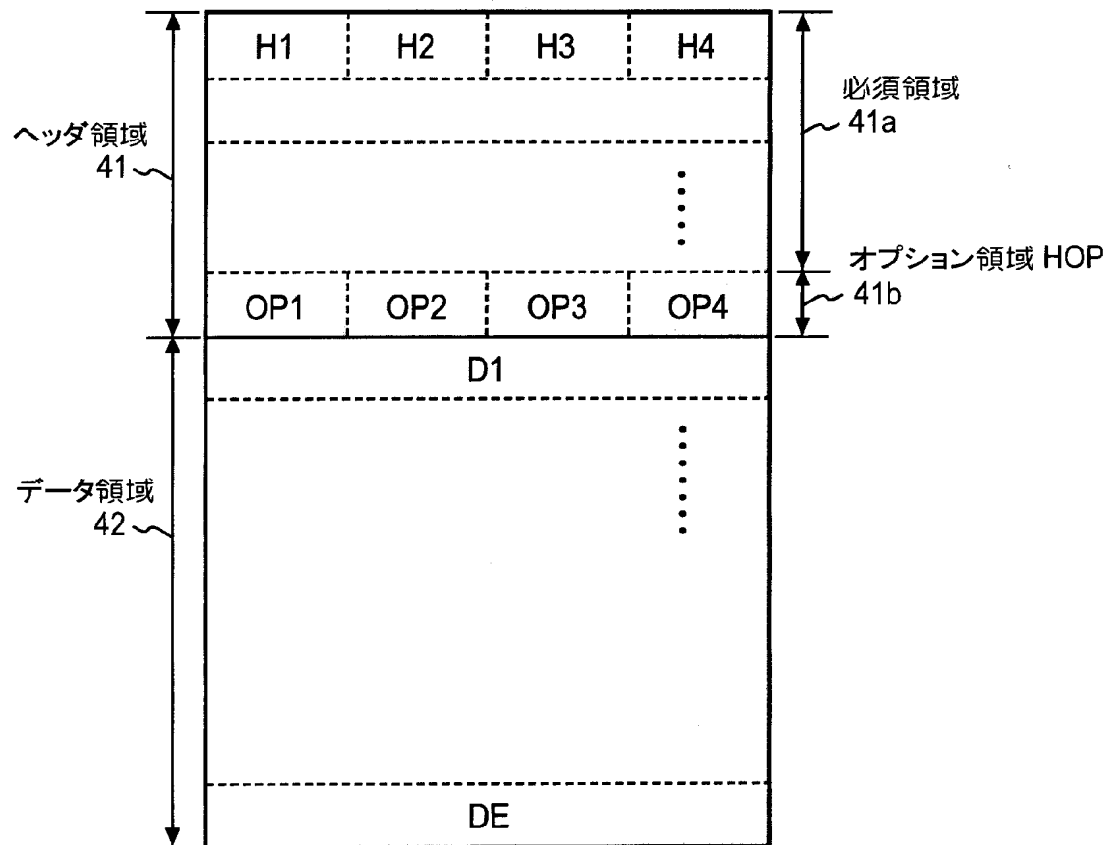


図11

[図12]

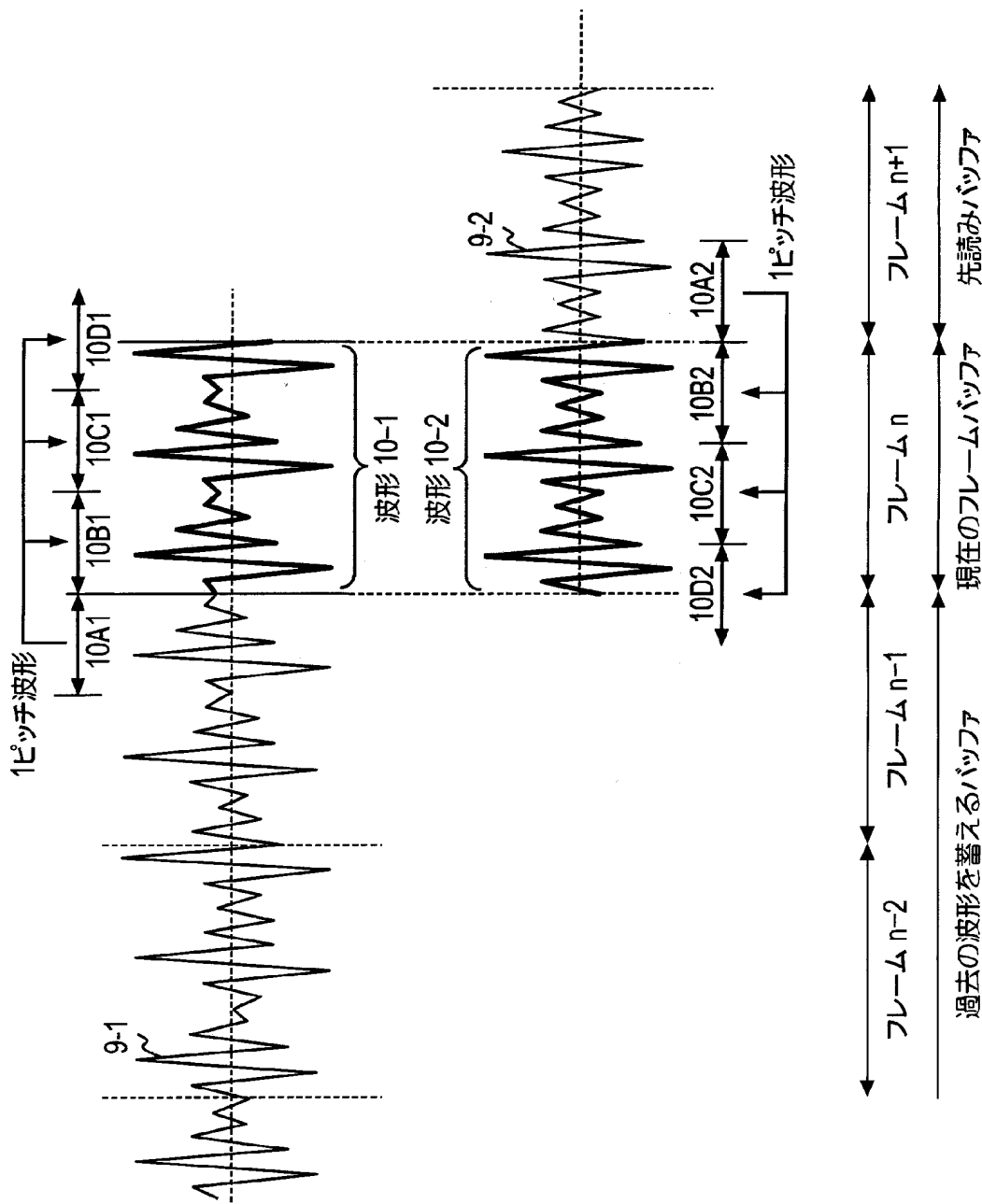


図12

[図13]

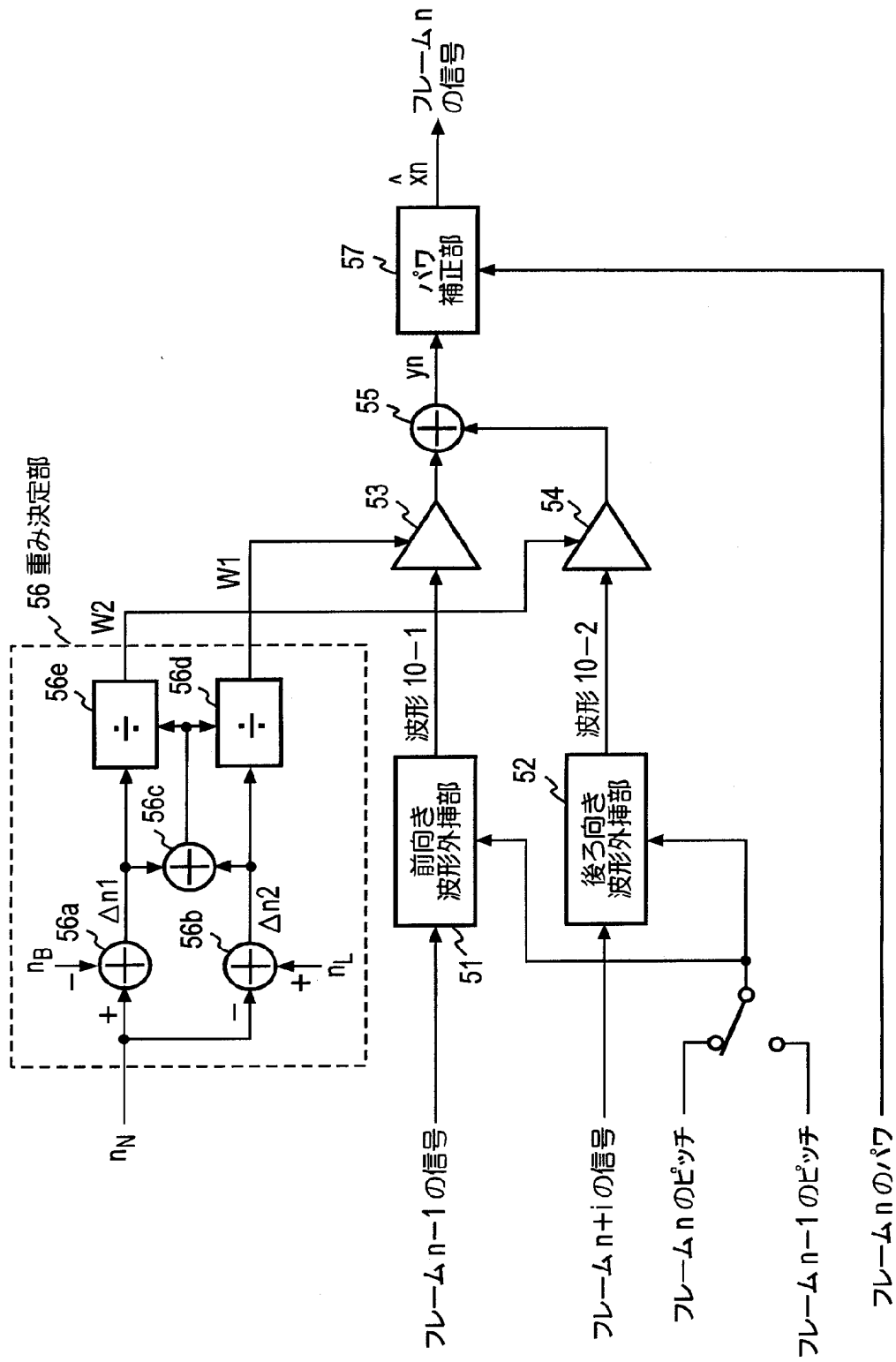


図13

[図14]

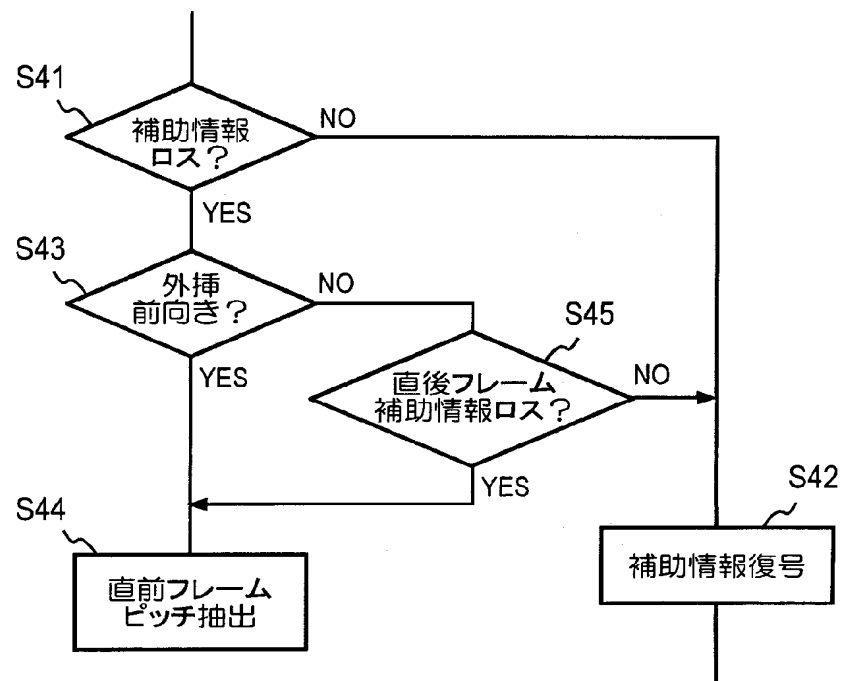


図14

[図15]

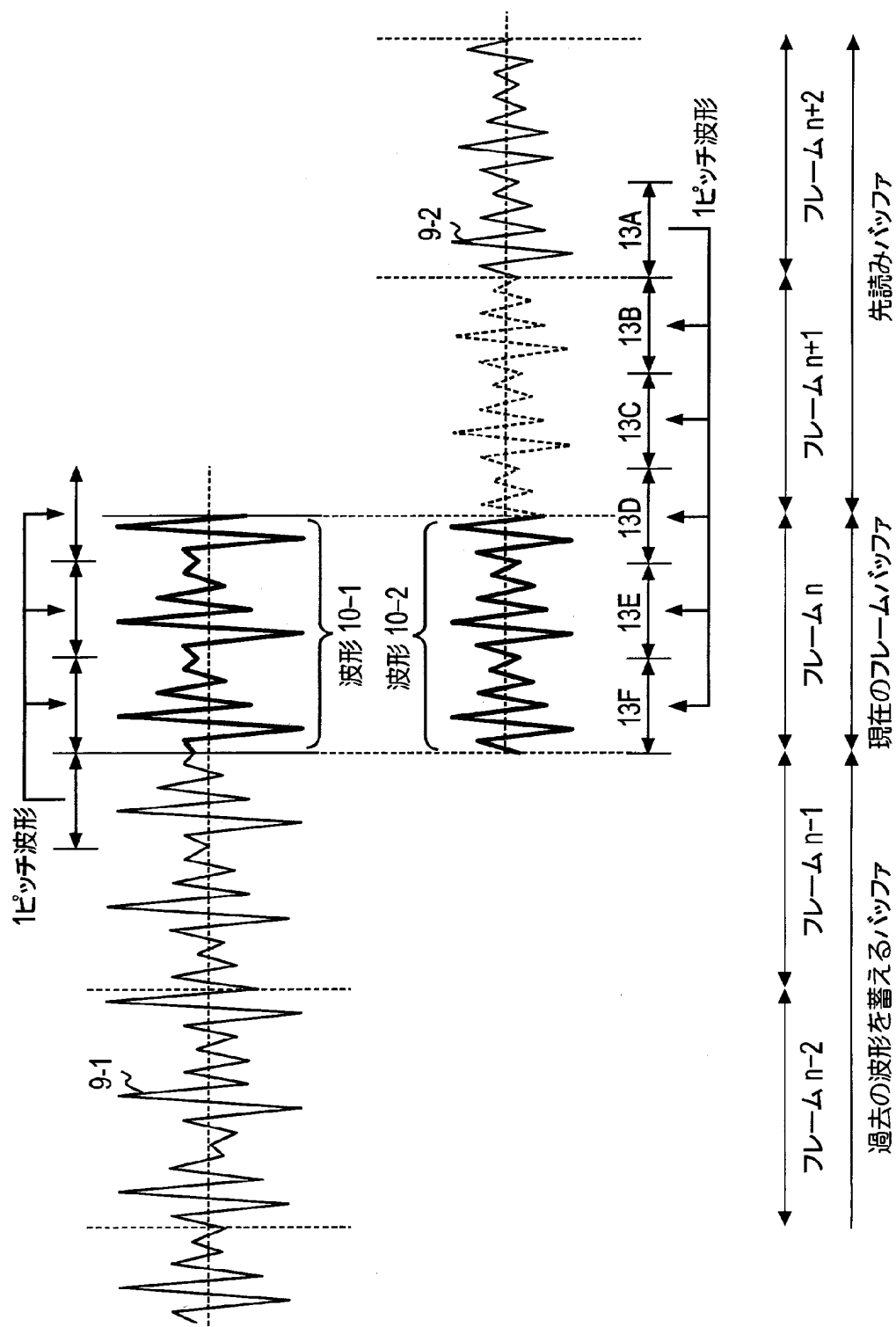


図15

[図16]

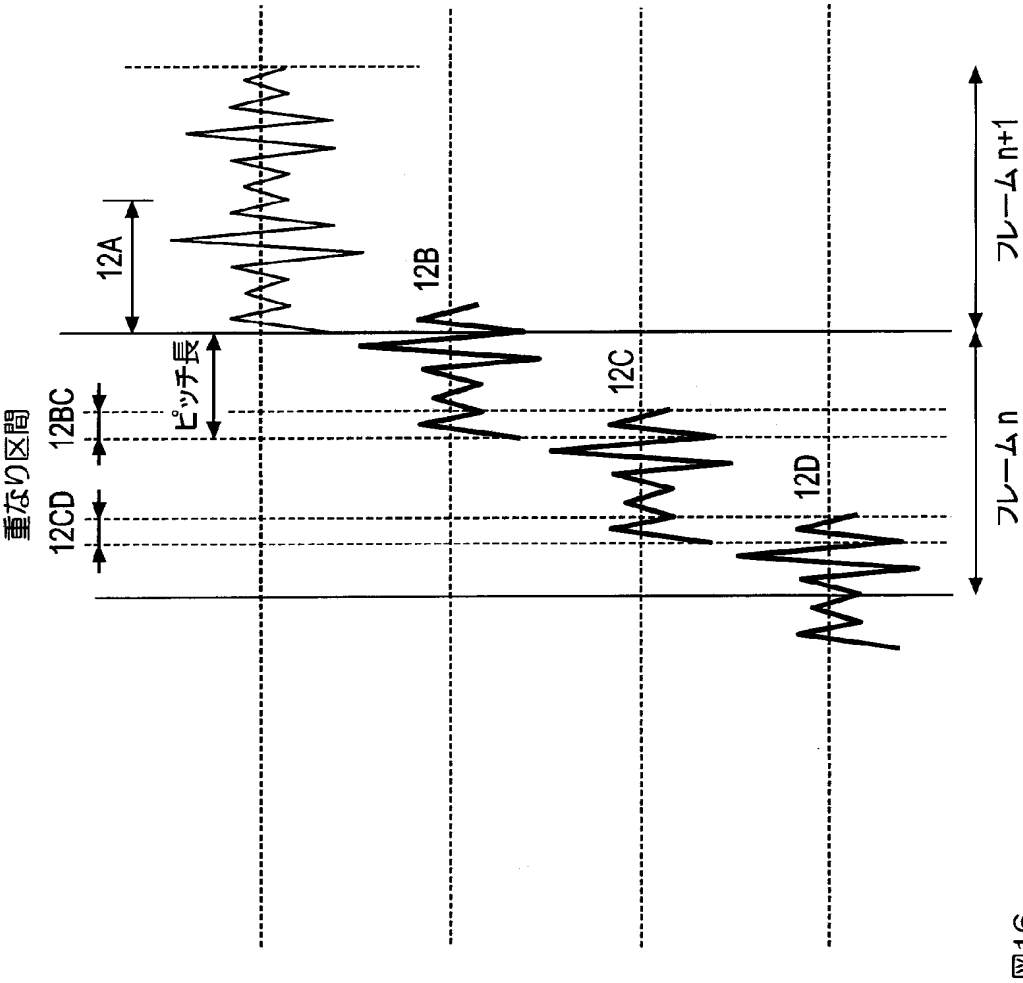


図16

[図17]

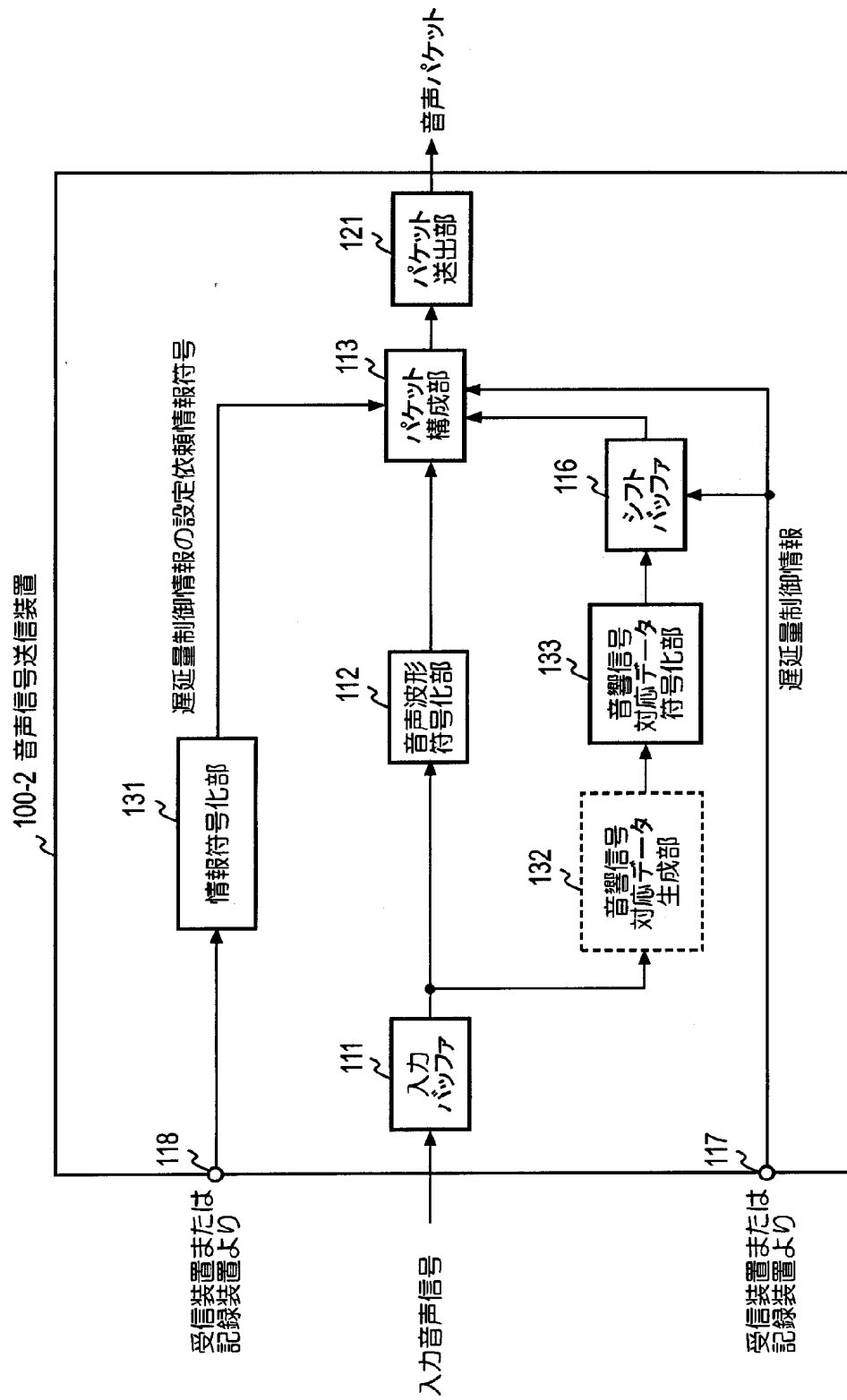


図17

[図18]

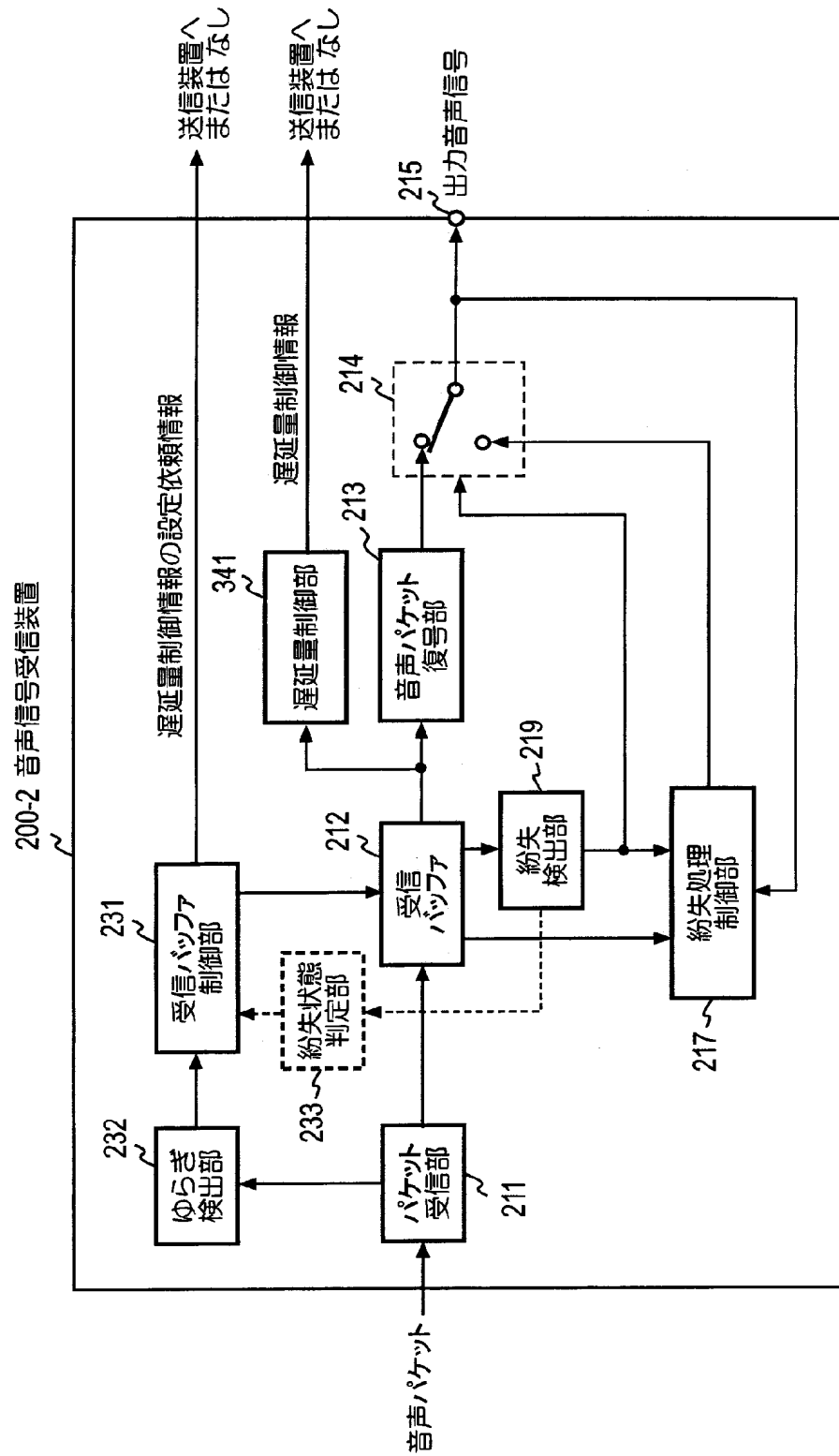


図18

[図19]

観測されたゆらぎ時間	バッファに蓄積すべき パケットの量 (蓄積パケット数の目標値)
15 ms まで	1 フレーム
32 ms まで	2 フレーム
48 ms まで	3 フレーム
64 ms まで	4 フレーム
80 ms まで	5 フレーム
96 ms まで	6 フレーム
128 ms まで	8 フレーム
160 ms まで	10 フレーム
192 ms まで	12 フレーム
240 ms まで	15 フレーム
320 ms まで	20 フレーム
400 ms まで	25 フレーム
480 ms まで	30 フレーム
640 ms まで	40 フレーム
800 ms まで	50 フレーム

図19

[図20]

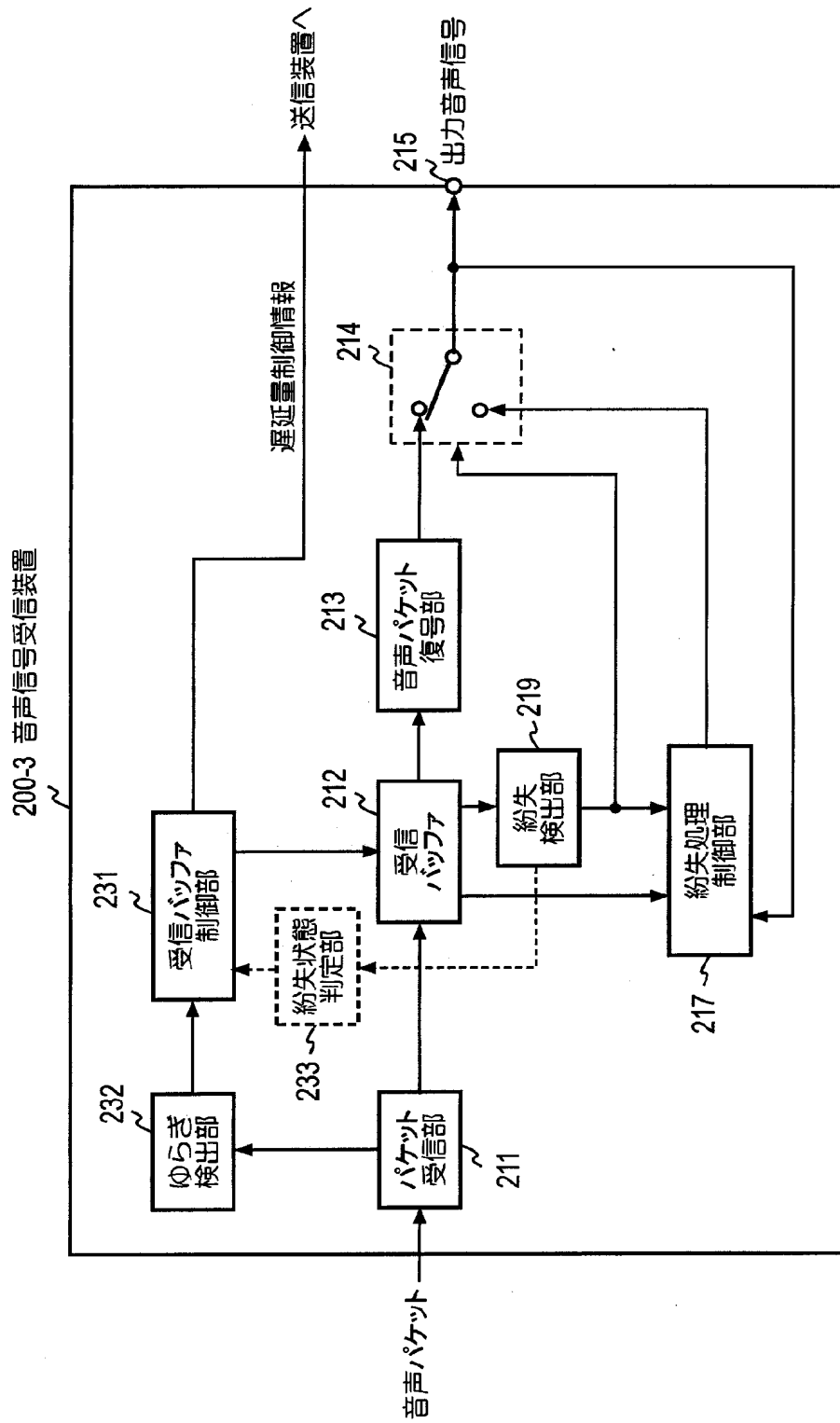
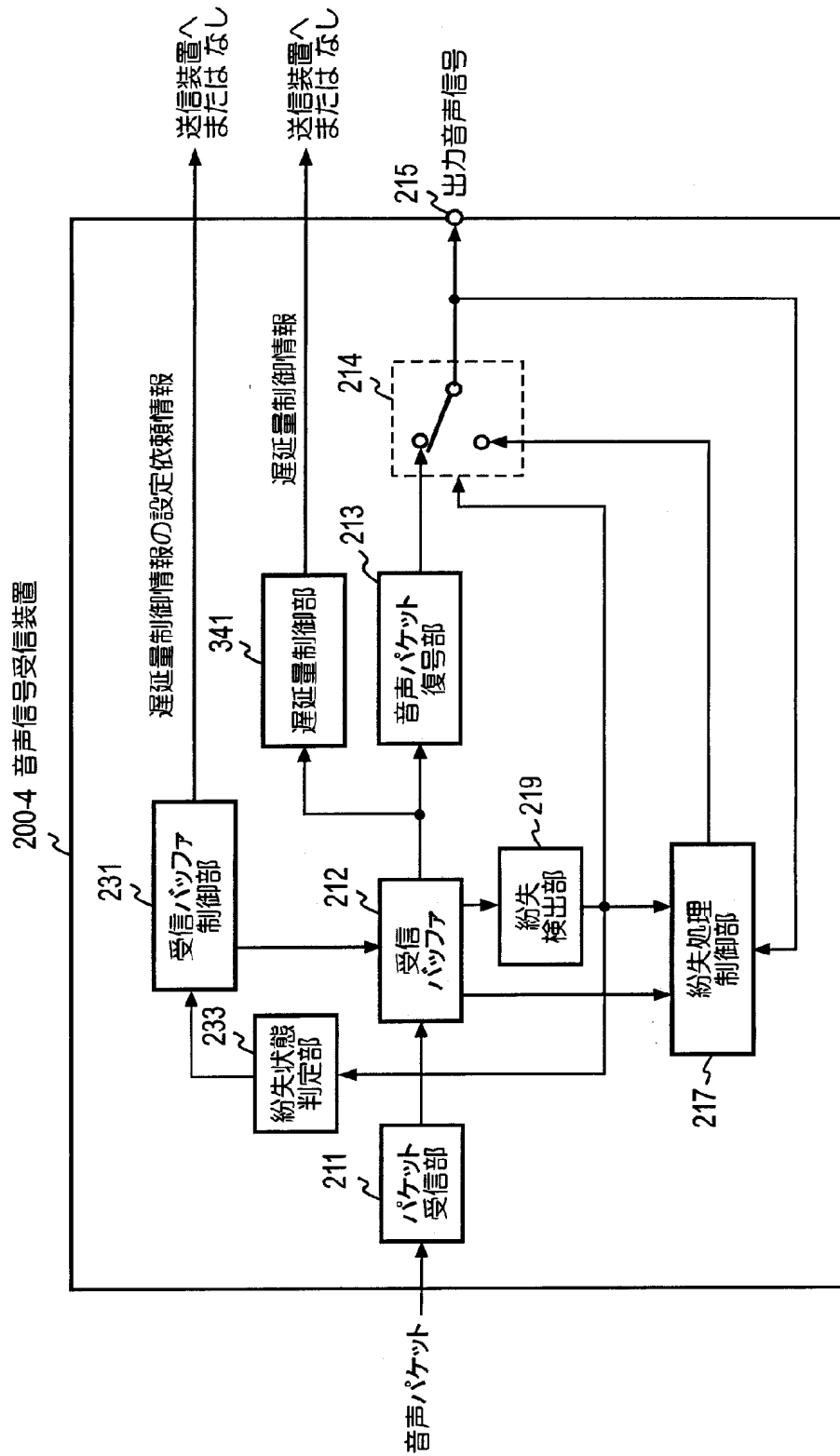


図20

[図21]



21

[図22]

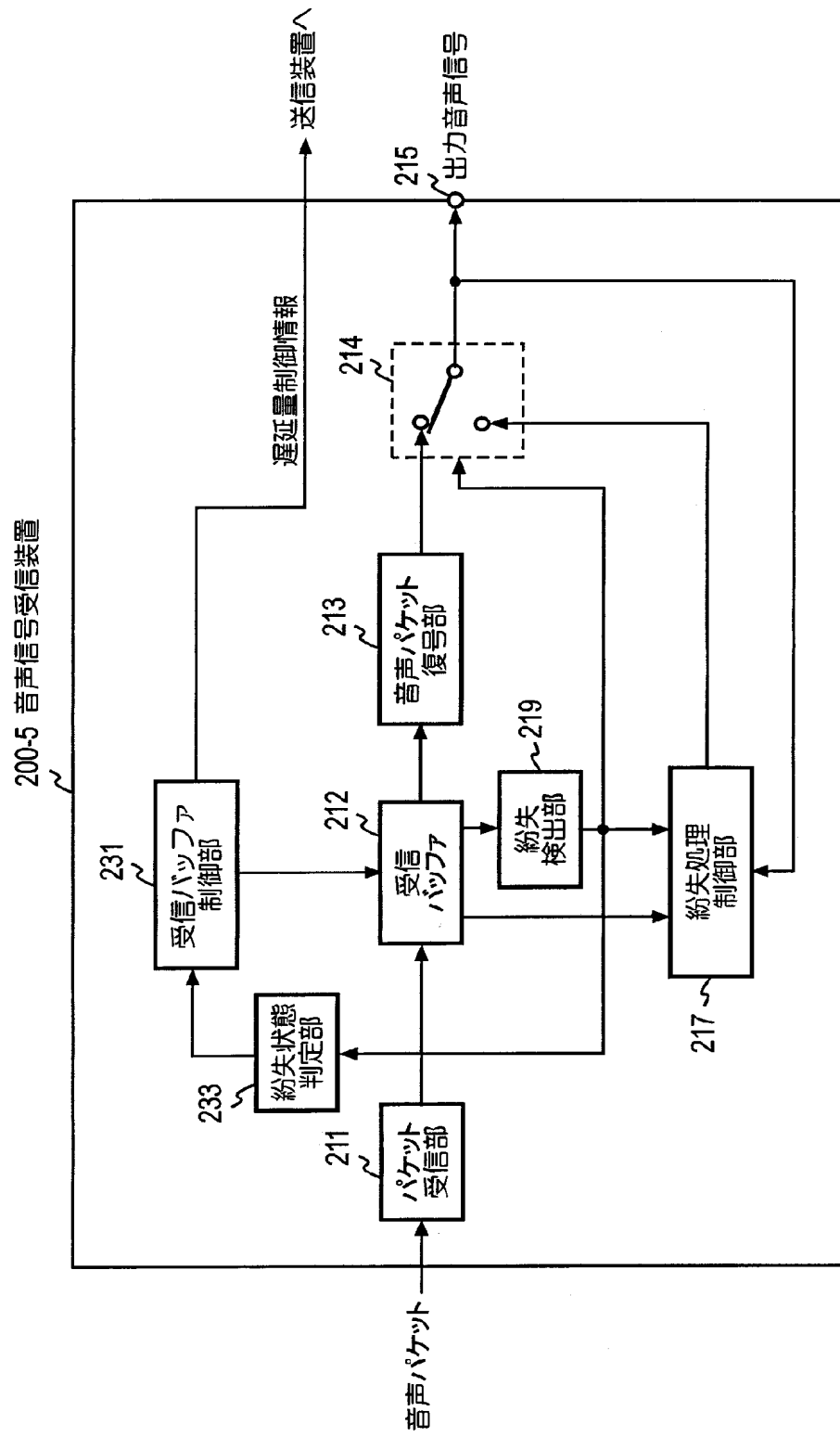


図22

[図23]

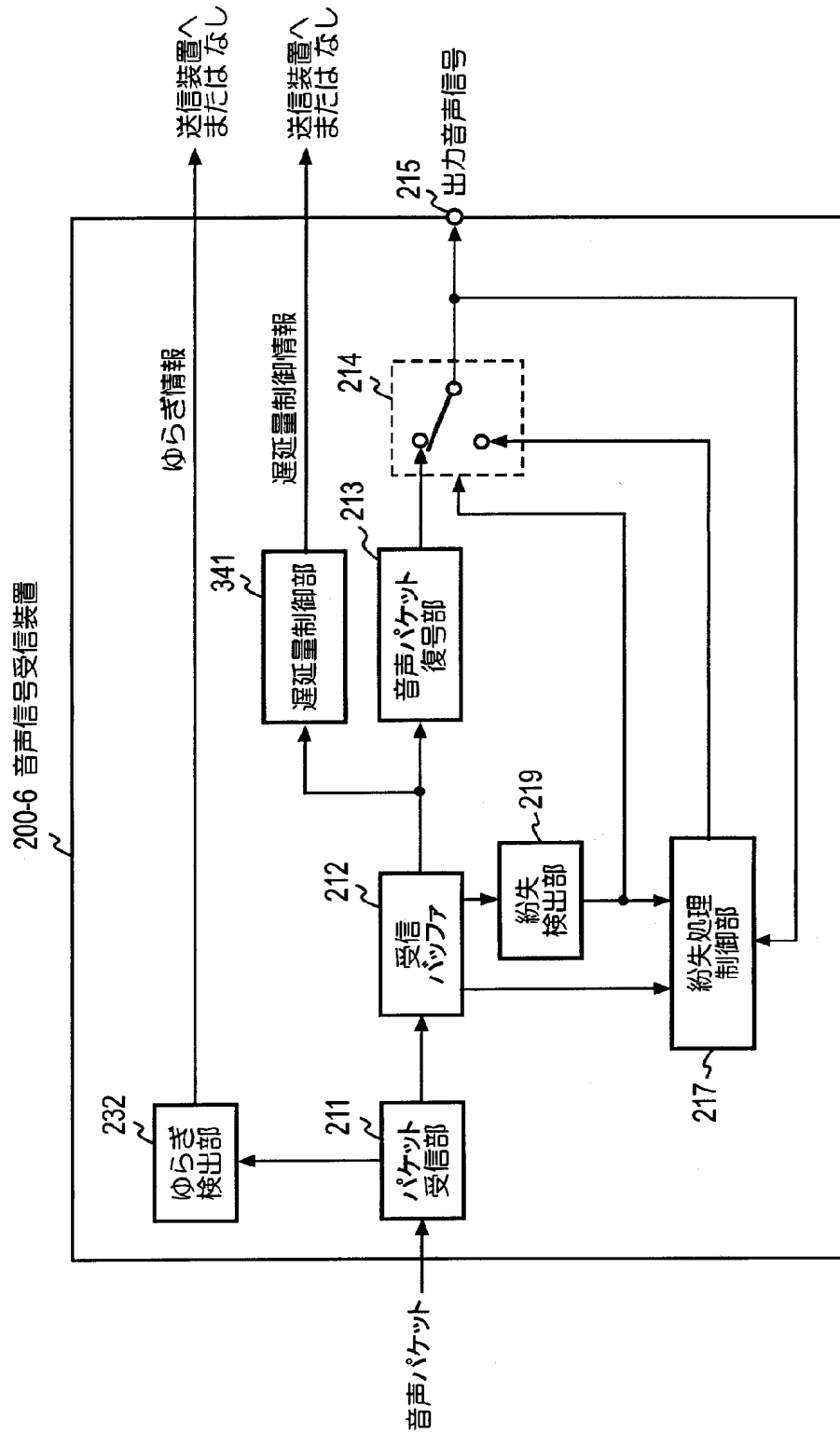
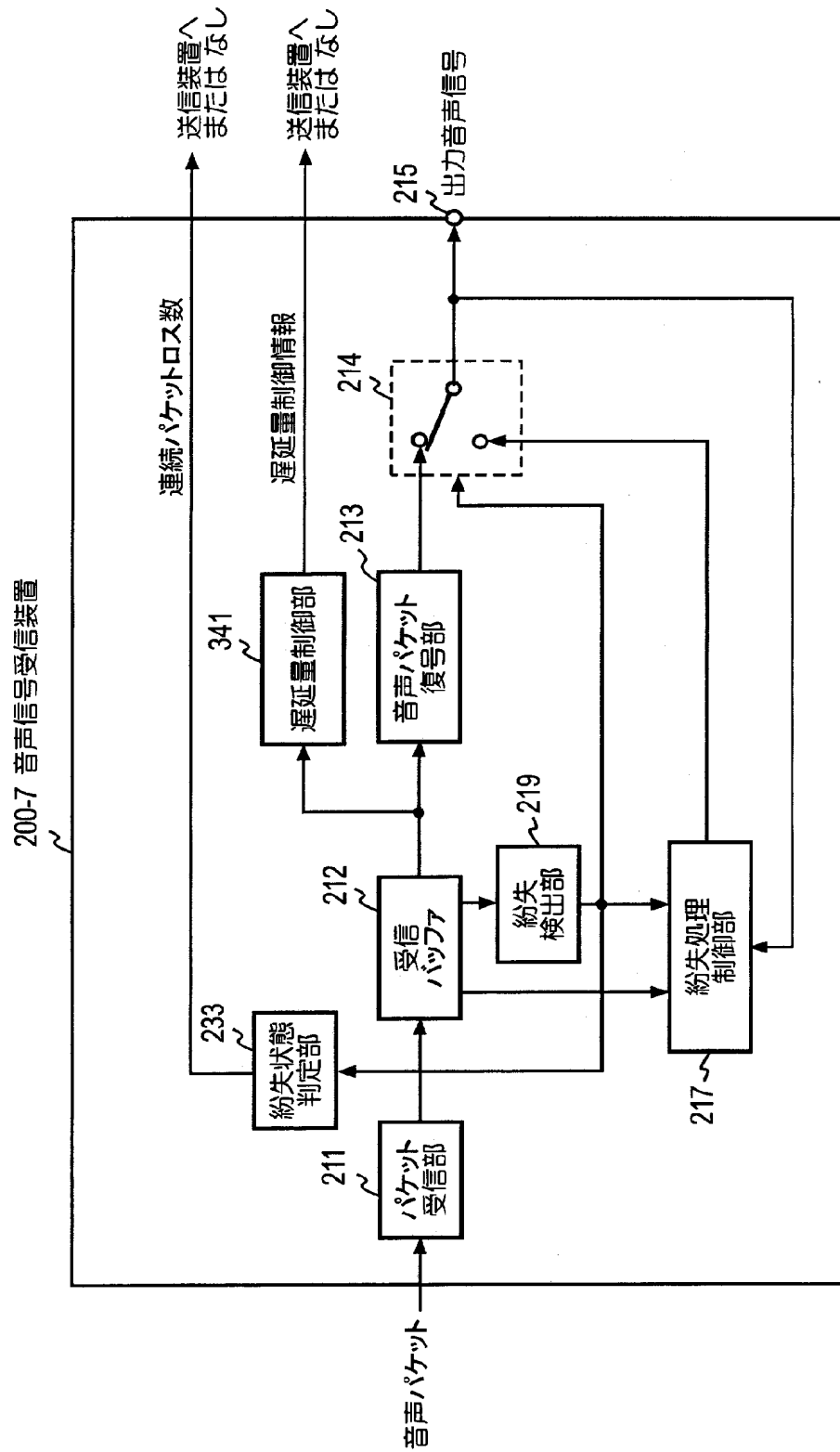


図23

[図24]



[図25]

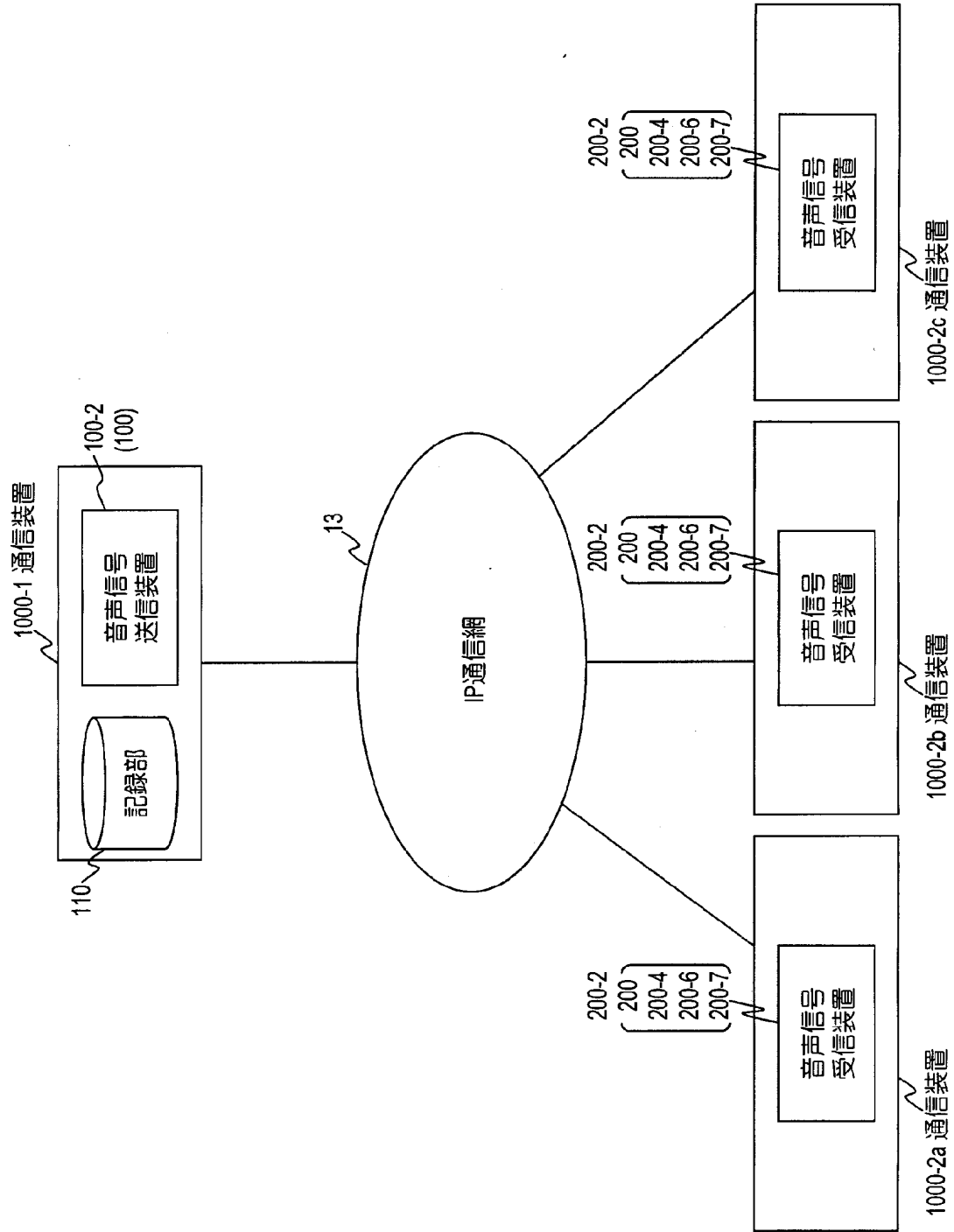


図25

[図26]

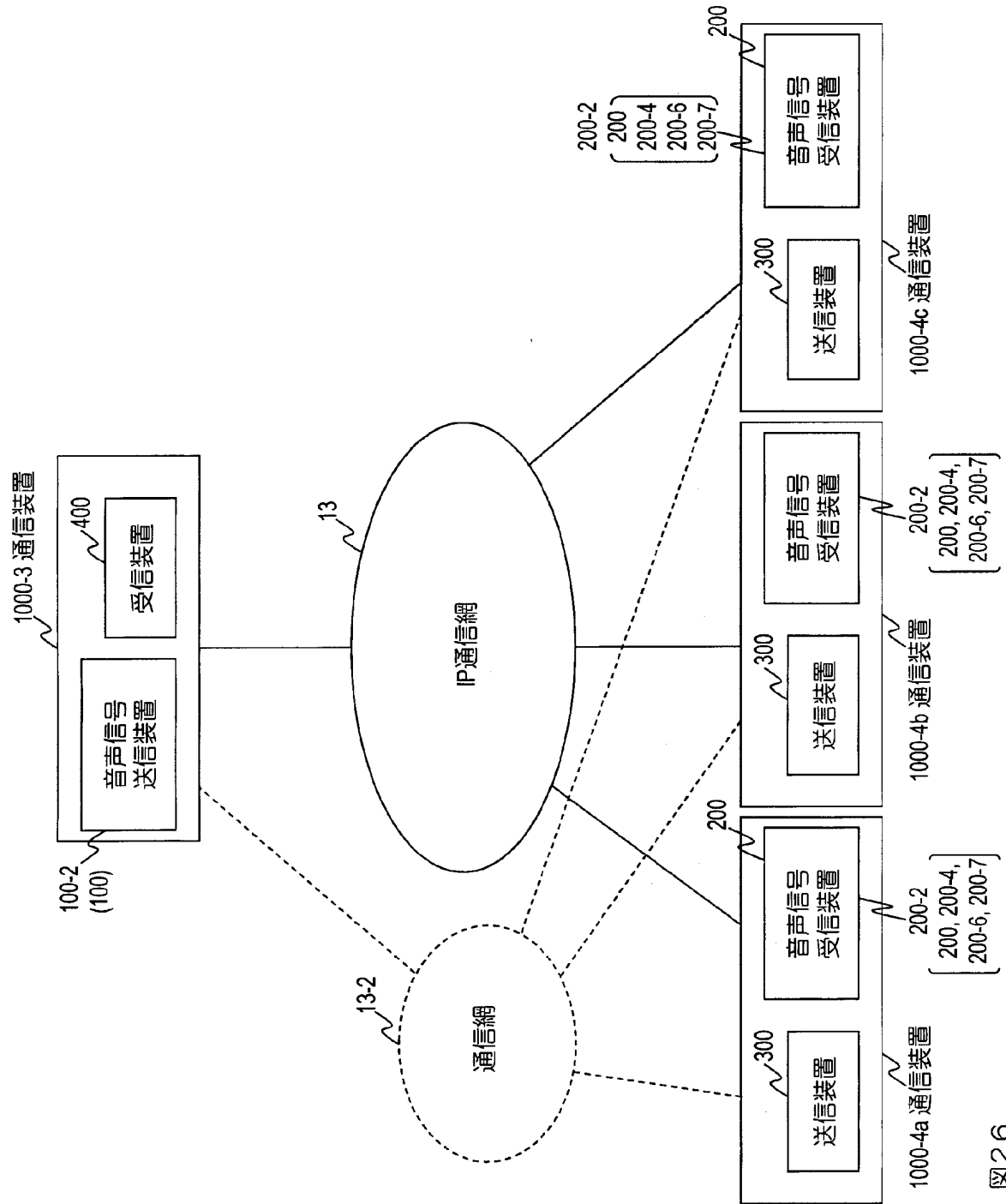


図26

[図27]

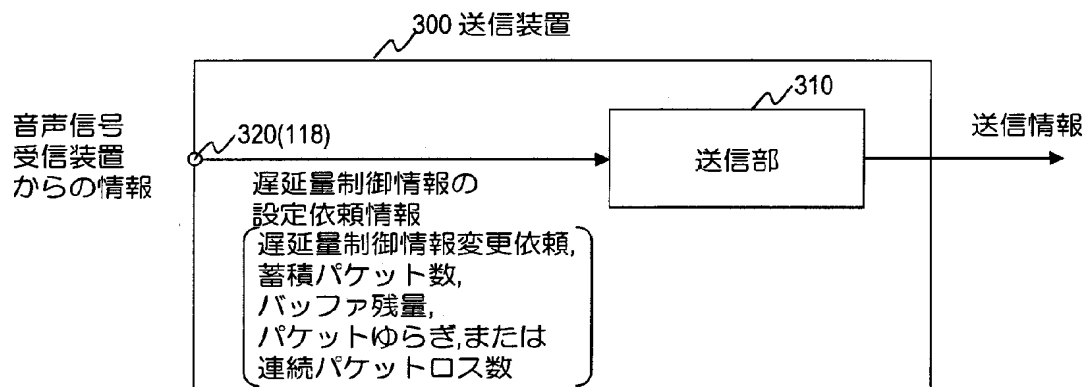


図27

[図28]

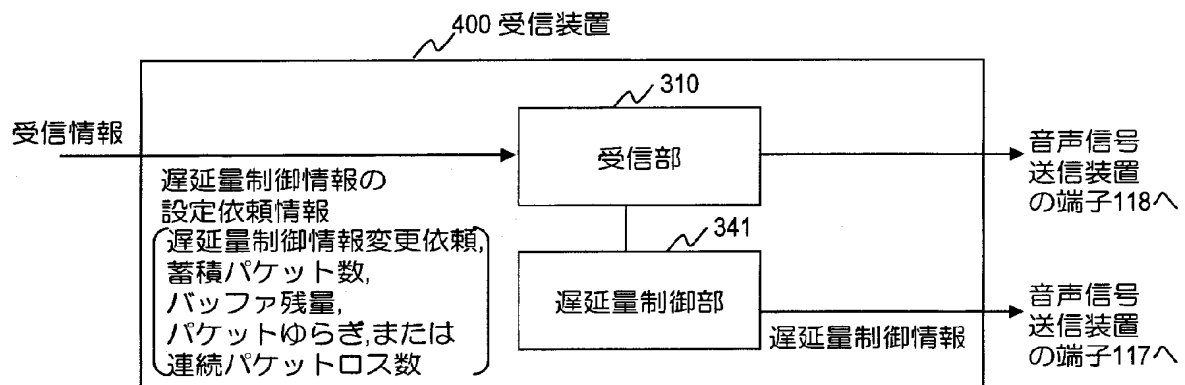


図28

[図29]

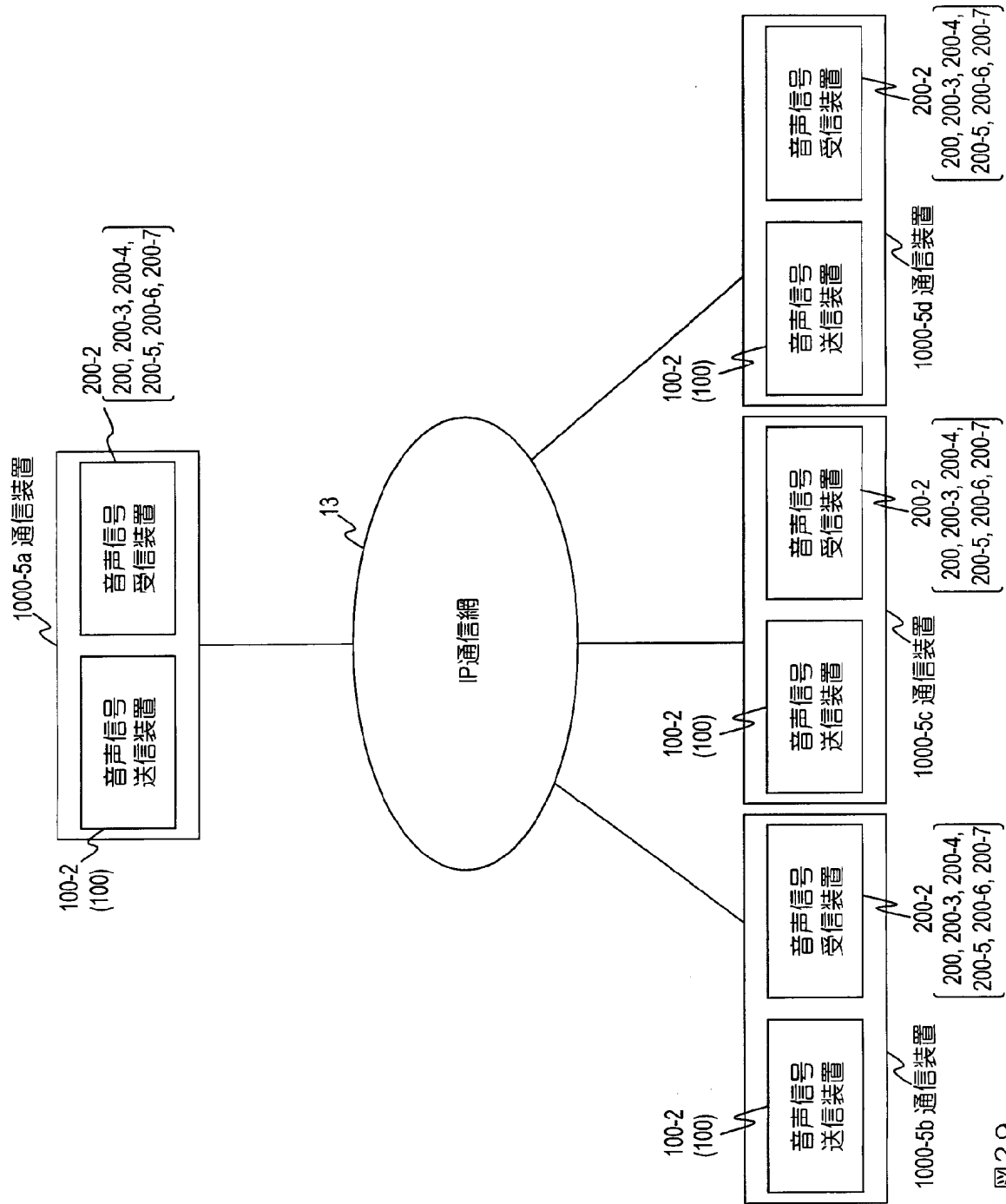
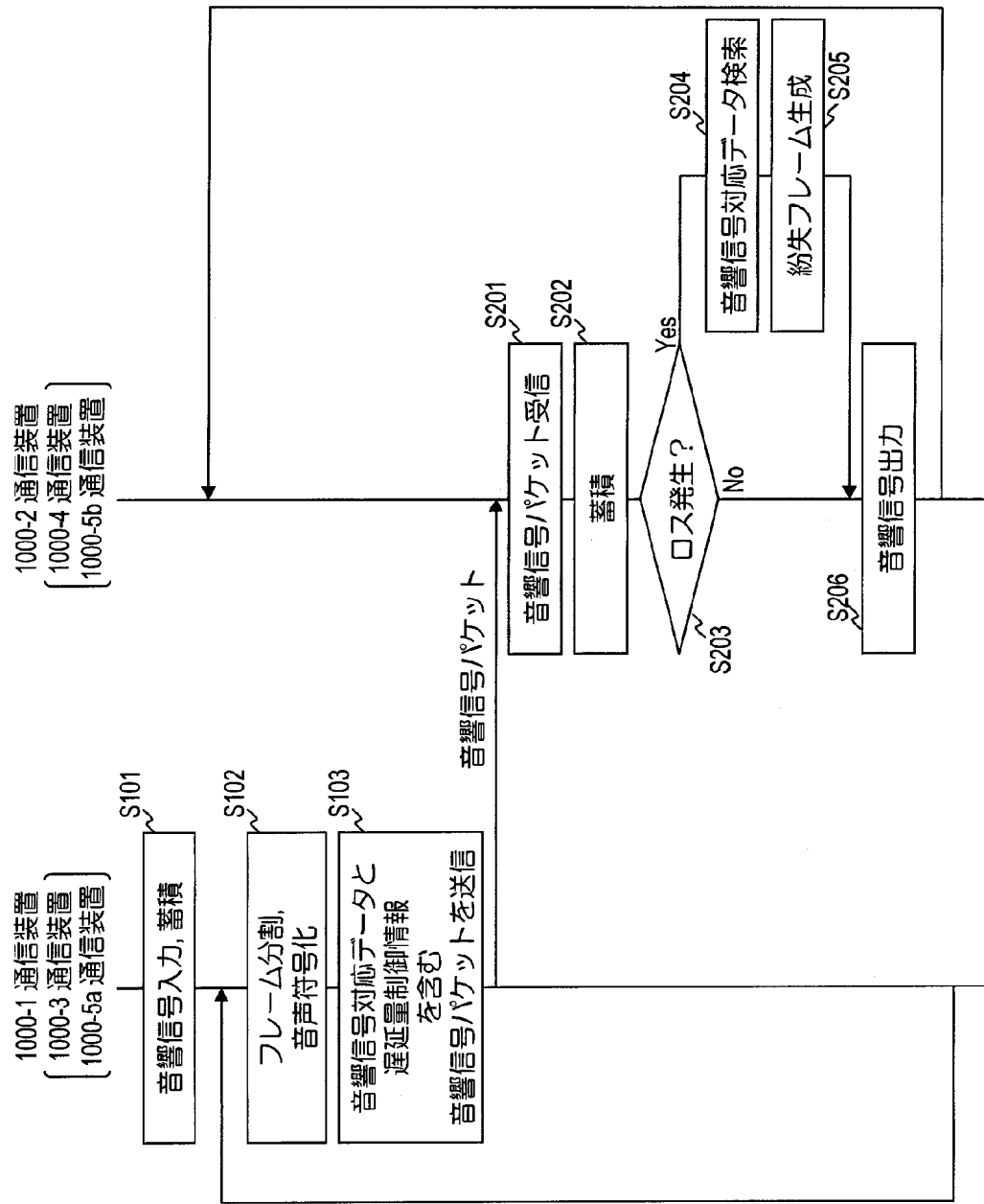


図29

[図30]



[図31]

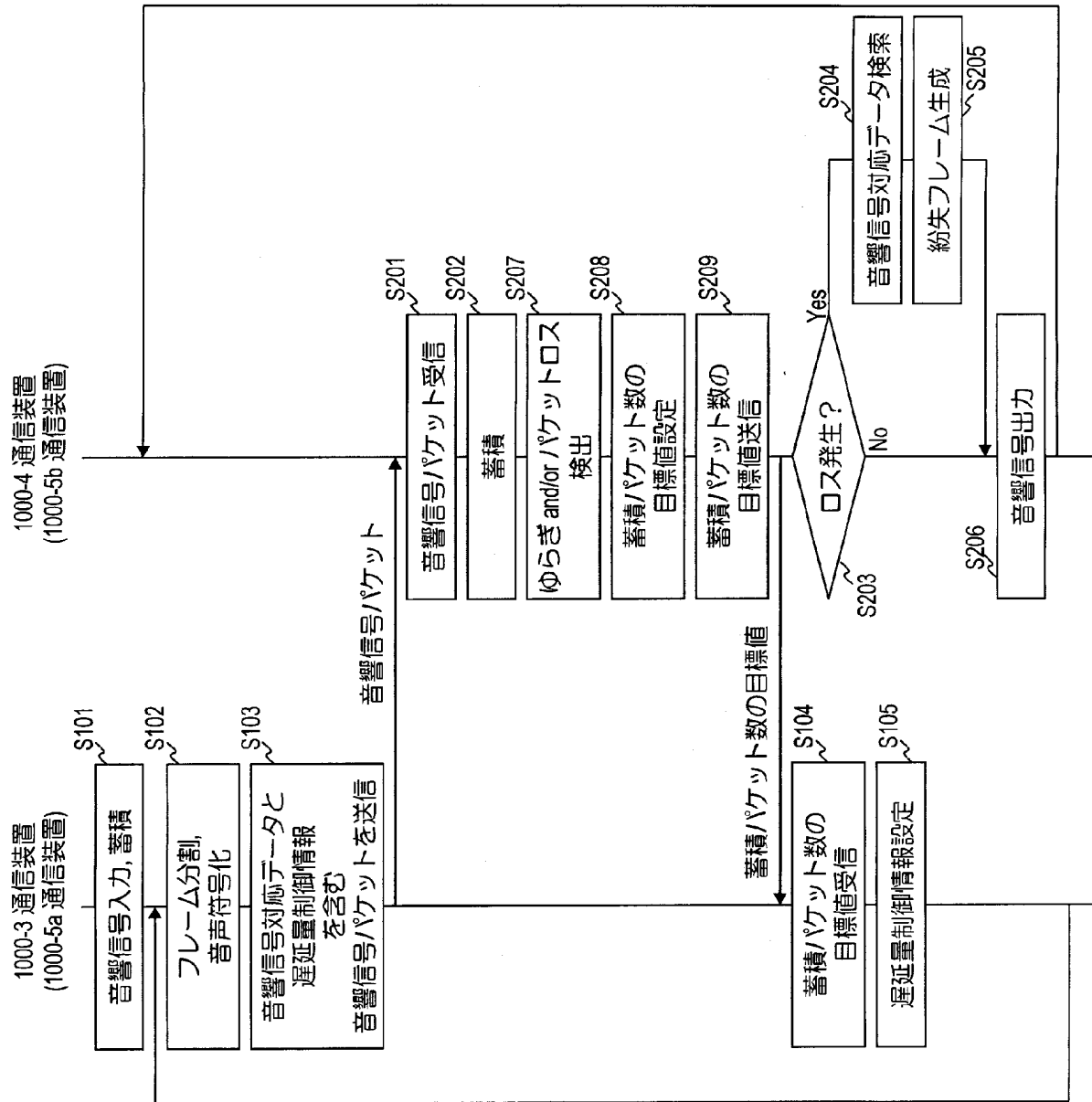


図31

[図32]

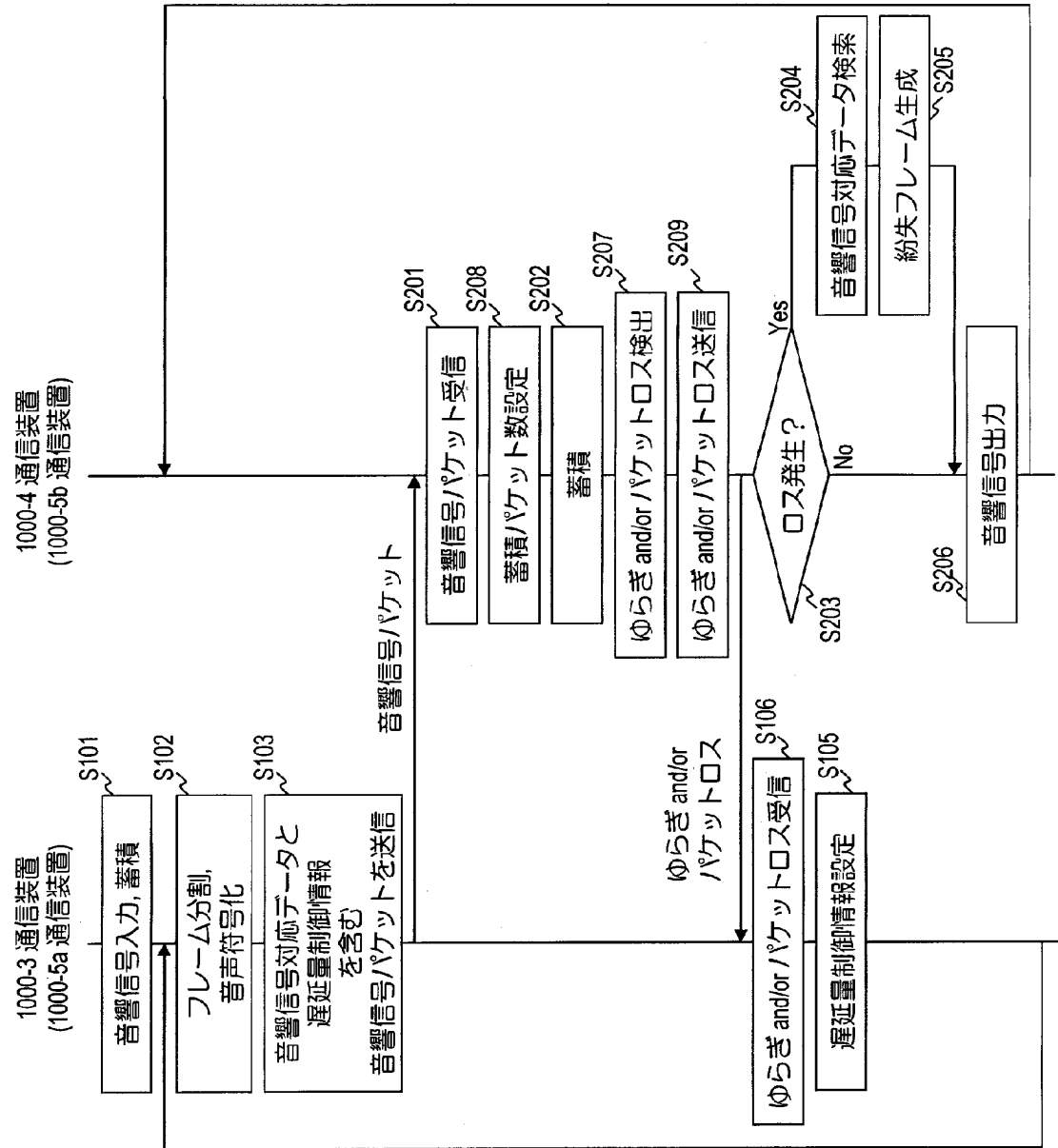


図32

[図33]

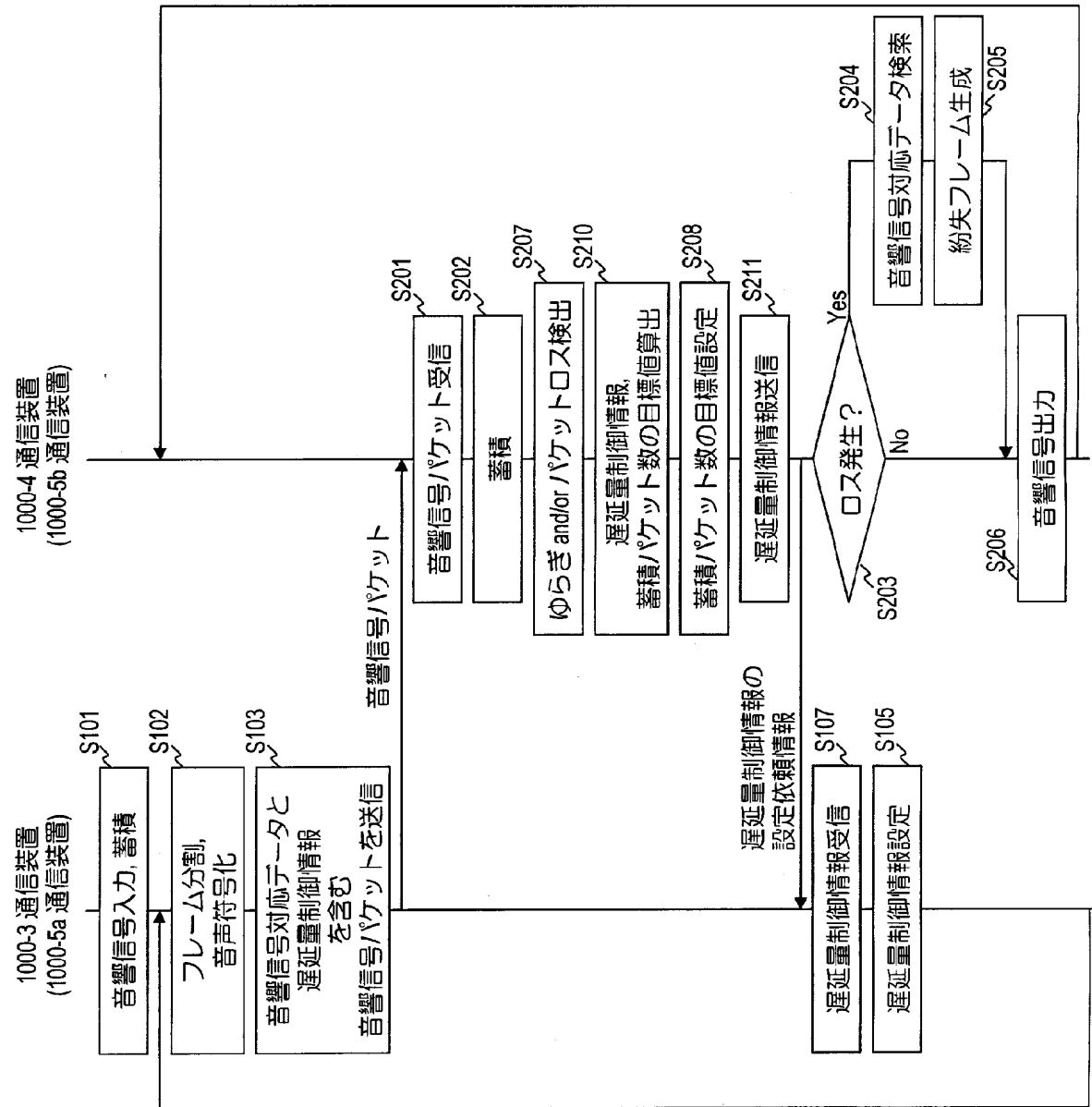


図33

[図34]

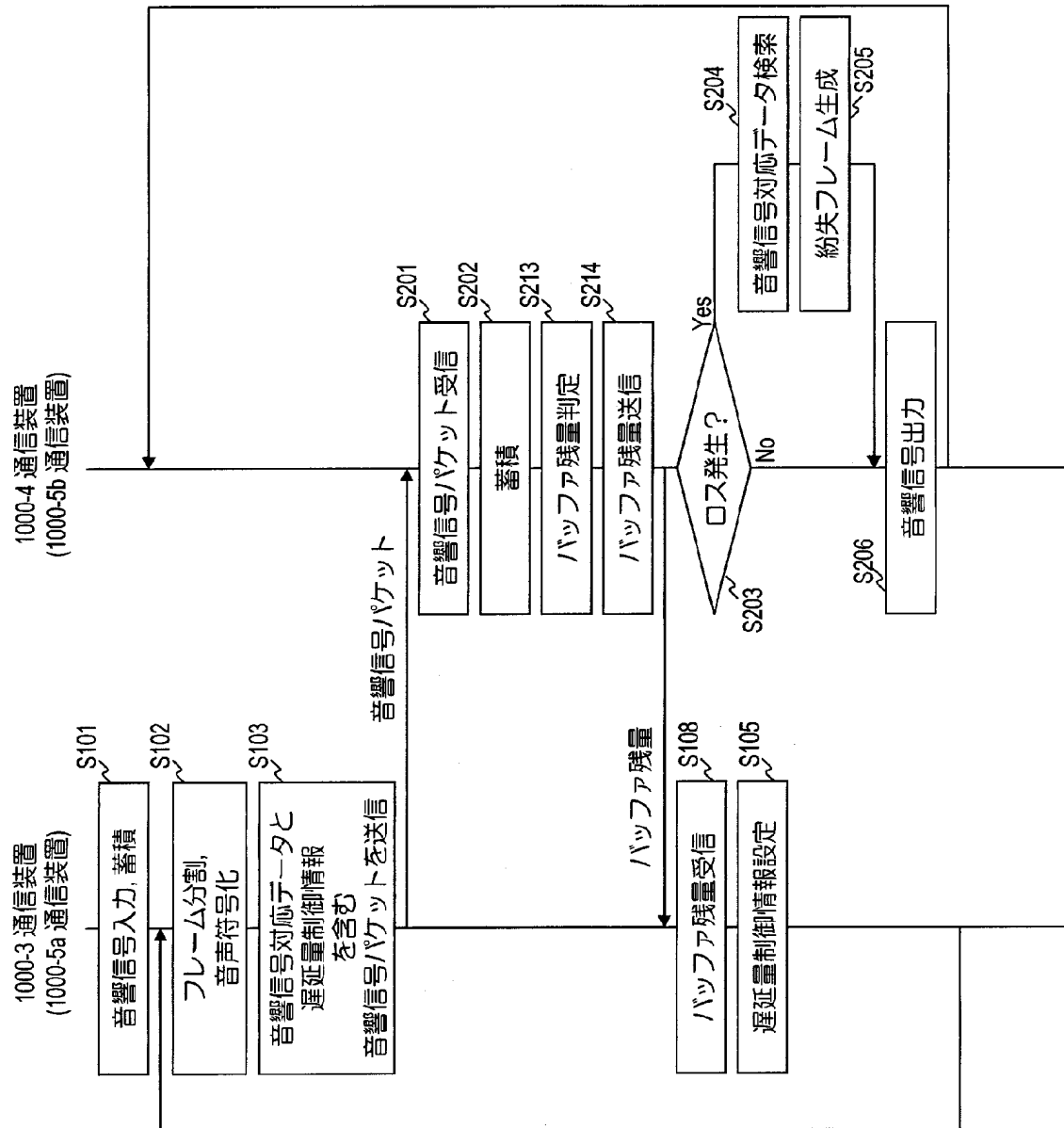


図34

[図35]

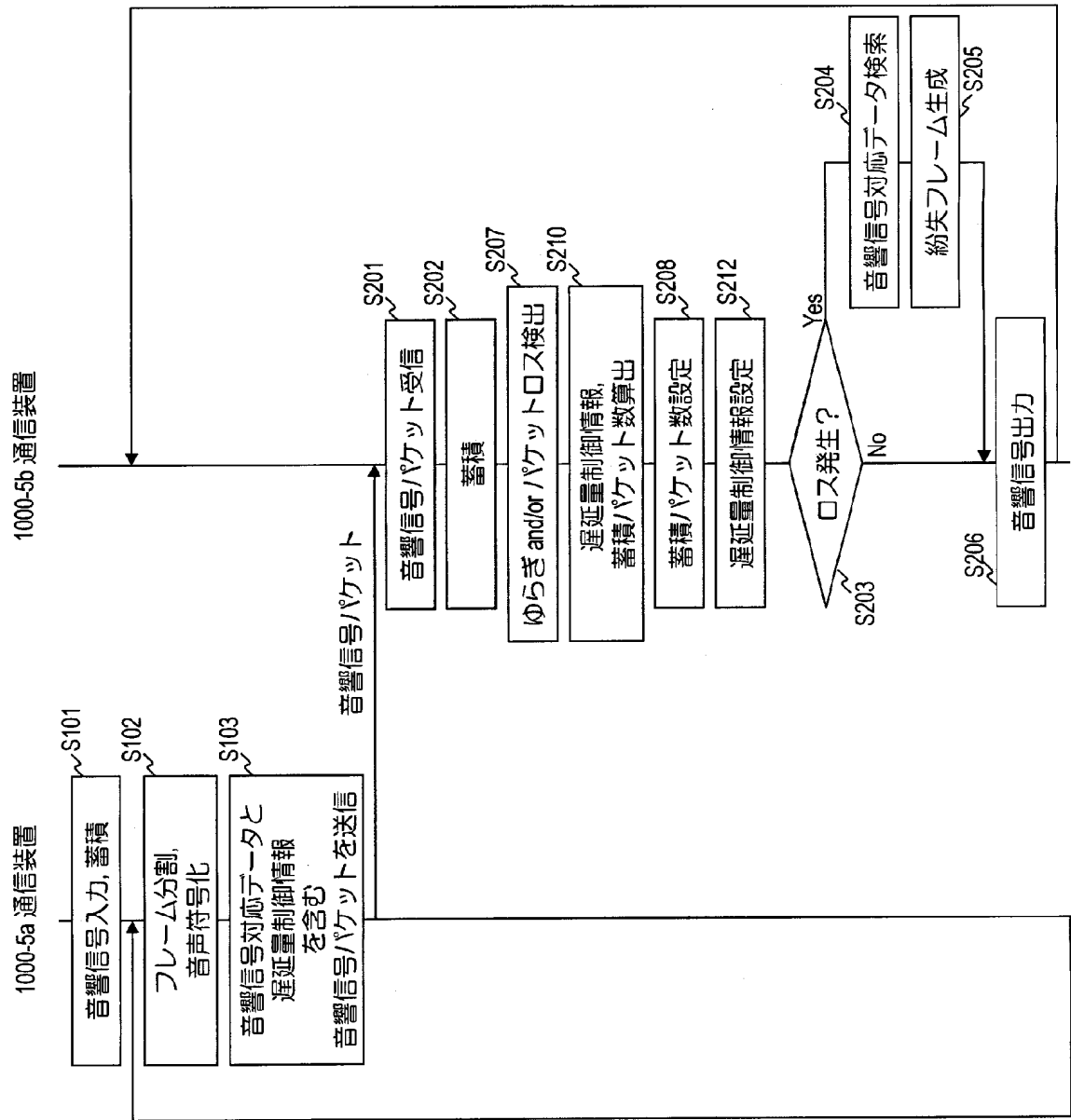


図35

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2005/008495

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl.⁷ G10L19/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl.⁷ G10L19/00-19/14, H04L1/00, 12/56

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2005
Kokai Jitsuyo Shinan Koho	1971-2005	Toroku Jitsuyo Shinan Koho	1994-2005

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

JSTPlus (JOIS)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP 2004-80625 A (Matsushita Electric Industrial Co., Ltd.), 11 March, 2004 (11.03.04), Full text; Figs. 1 to 8 (Family: none)	1-19
A	JP 2003-50598 A (Mitsubishi Electric Corp.), 21 February, 2003 (21.02.03), Full text; Figs. 1 to 27 (Family: none)	1-19
A	JP 2003-316391 A (NEC Corp.), 07 November, 2003 (07.11.03), Full text; Figs. 1 to 12 & US 2003/0200083 A1	1-19



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T"

later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X"

document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y"

document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&"

document member of the same patent family

Date of the actual completion of the international search

08 August, 2005 (08.08.05)

Date of mailing of the international search report

23 August, 2005 (23.08.05)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2005/008495

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2002-515612 A (Telefonaktiebolaget LM Ericsson), 28 May, 2002 (28.05.02), Full text; Figs. 1 to 10 & WO 1999/059282 A2 & EP 1076965 A2 & US 6721327 B1	1-19
A	JP 2004-120619 A (KDDI Corp.), 15 April, 2004 (15.04.04), Full text; Figs. 1 to 19 (Family: none)	1-19
A	JP 2002-268697 A (NEC Corp.), 20 September, 2002 (20.09.02), Full text; Figs. 1 to 8 & EP 1241664 A2 & US 2002/0169859 A1	1-19

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int.Cl.⁷ G10L19/00

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int.Cl.⁷ G10L19/00-19/14, H04L1/00, 12/56

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2005年
日本国実用新案登録公報	1996-2005年
日本国登録実用新案公報	1994-2005年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

JSTPlus (JOIS)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
X	JP 2004-80625 A (松下電器産業株式会社) 2004.03.11, 全文, 図1-8 (ファミリーなし)	1-19
A	JP 2003-50598 A (三菱電機株式会社) 2003.02.21, 全文, 図1-27 (ファミリーなし)	1-19
A	JP 2003-316391 A (日本電気株式会社) 2003.11.07, 全文, 図1-12 & US 2003/0200083 A1	1-19

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」 特に関連のある文献ではなく、一般的技術水準を示すもの
「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
「O」 口頭による開示、使用、展示等に言及する文献
「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「&」 同一パテントファミリー文献

国際調査を完了した日

08.08.2005

国際調査報告の発送日

23.8.2005

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

山下 剛史

電話番号 03-3581-1101 内線 3541

5Z

8946

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	JP 2002-515612 A (テレフォンアクチーボラゲット エル エム エ リクソン), 2002. 05. 28, 全文, 図 1-10 & WO 1999/059282 A2 & EP 1076965 A2 & US 6721327 B1	1-19
A	JP 2004-120619 A (KDD I 株式会社) 2004. 04. 15, 全文, 図 1-19 (ファミリーなし)	1-19
A	JP 2002-268697 A (日本電気株式会社) 2002. 09. 20, 全文, 図 1-8 & EP 1241664 A2 & US 2002/0169859 A1	1-19